Access DB# 53877

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____

Art Unit: _____ Phone Number 30_____ Serial Number: _____

Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

_____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: Beverly C4994

Searcher Phone #: _____

Searcher Location: _____

Date Searcher Picked Up: _____

Date Completed: 10-31-01

Searcher Prep & Review Time: _____

Clerical Prep Time: _____

Online Time: _____

Type of Search

NA Sequence (#)_____

AA Sequence (#)_____

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN _____

Dialog _____

Questel/Orbit _____

Dr. Links _____

Lexis/Nexis _____

Sequence Systems _____

WWW/Internet _____

Other (specify) CGN

BEST AVAILABLE COPY

This Page Blank (uspto)

[redacted]

STIC-Biotech/ChemLib 53877

From: Chan, Christina
Sent: Monday, October 29, 2001 10:40 AM
To: Chunduru, Suryaprahbha; STIC-Biotech/ChemLib
Subject: RE: ref to rush seq search

Please rush. Thanks Chris

-----Original Message-----
From: Chunduru, Suryaprahbha
Sent: Friday, October 26, 2001 10:23 AM
T : Chan, Christina
Subject: ref to rush seq search

Hi Christina,
I need a rush sequence search on the following. I need your approval on this rush sequence search since the case is due for this biweek.

1. Application Serial NO. 09/619049. SEQ ID NO. 855 and 854.
(Search is requested for all commercial, nucleic acid, oligomer, pending and issued patent databases).

Thank you
Suryaprabha Chunduru
Examiner
Art Unit 1656
Room No. 8B13
Mail Room No. 10E12.
Tel. No. 305-1004.

Point of Contact:
Beverly Shears
Technical Info. Specialist
CM1 2C14 Tel: 308-4994
1E05

RECEIVED
OCT 29 2001
STIC

This Page Blank (uspto)

```
Db    541 gttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaactttgat 600

Qy    601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660

Qy    661 ggatccttagagattctaattataacaatgatatgctttgtcctaacaatcgcagtatat 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ggatccttagagattctaattataacaatgatatgctttgtcctaacaatcgcagtatat 720

Qy    721 ctatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcg 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ctatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcg 780

Qy    781 aggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgc 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 aggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgc 840

Qy    841 tctccagctggttacagctcgcacttttccggatggcttccaacctctggctctccgtc 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D     841 tctccagctggttacagctcgcacttttccggatggcttccaacctctggctctccgtc 900

Qy    901 atcagctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaactatcgg 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 atcagctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaactatcgg 960

Qy    961 ttcctgcggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttat 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ttcctgcggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttat 1020

Qy   1021 atagttaatcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 atagttaatcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggt 1080

Qy   1081 tttattcggtgctcggtcaaagattggcacccatccgtctggatctatataagtggaccg 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 tttattcggtgctcggtcaaagattggcacccatccgtctggatctatataagtggaccg 1140

Qy   1141 tcgctggccctgagcactttcaatgtcgccatgttcgccctgacagccatttacattagg 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 tcgctggccctgagcactttcaatgtcgccatgttcgccctgacagccatttacattagg 1200

Qy   1201 aaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggattaactgcata 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 aaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggattaactgcata 1260

Qy   1261 aactttgacagccagacgtaagatatgataaaaccttggaccttggtaaactagttatcc 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D     261 aactttgacagccagacgtaagatatgataaaaccttggaccttggtaaactagttatcc 1320

Qy   1321 ttgccgtattatccacatagttacctacagttcctgcggctctccatcgtgatgggcctt 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 ttgccgtattatccacatagttacctacagttcctgcggctctccatcgtgatgggcctt 1380

Qy   1381 acttggatattcaatgtcattccgtattctgcacggctccacattttctgggaatgggtc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 acttggatattcaatgtcattccgtattctgcacggctccacattttctgggaatgggtc 1440

Qy   1441 gggataatatccgagtattttcacagcgcgtttggaattgttttgttcgttctgctcgtc 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 gggataatatccgagtattttcacagcgcgtttggaattgttttgttcgttctgctcgtc 1500

Qy   1501 ctgaagcgcagcacatggactctaatgatggattct 1536
          |||||||||||||||||||||||||||||||||||
Db   1501 ctgaagcgcagcacatggactctaatgatggattct 1536


RESULT   7
US-60-191-637-16753
; Sequence 16753, Application US/60191637
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
;  TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
;  TITLE OF INVENTION: USES THEREOF
;  FILE REFERENCE: CL000392
;  CURRENT APPLICATION NUMBER: US/60/191,637
;  CURRENT FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 42660
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16753
;   LENGTH: 1536
;   TYPE: DNA
;   ORGANISM: DROSOPHILA
US-60-191-637-16753


  Query Match             100.0%;  Score 1536;  DB 51;  Length 1536;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1536;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 atgaggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgaggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactca 60

Qy     61 aatgcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaagg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 aatgcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaagg 120

Qy    121 caaaacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactac 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 caaaacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactac 180

Qy    181 gagttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgt 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gagttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgt 240

Qy    241 gtttgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttggccaac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gtttgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttggccaac 300

Qy    301 ggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttc 360

Qy    361 acatacatggacctacaggcacgagtaccactcaccgatatggctattatcagagatgag 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 acatacatggacctacaggcacgagtaccactcaccgatatggctattatcagagatgag 420

Qy    421 ttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggtt 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggtt 480

Qy    481 agcattcaaatcttcaataaatgtggacttatagtttgyttccaggatggaaaattttgg 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 agcattcaaatcttcaataaatgtggacttatagtttygttccaggatggaaaattttgg 540

Qy    541 gttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaactttgat 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 gttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaactttgat 600

Qy    601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660

Qy    661 ggatccttagagattctaattataacaatgatatgctttgtcctaacaatcgcagtatat 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ggatccttagagattctaattataacaatgatatgctttgtcctaacaatcgcagtatat 720

Qy    721 ctatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcg 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ctatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcg 780

Qy    781 aggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgc 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 aggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgc 840
```

This Page Blank (uspto)

```
Qy    841 tctccagctggttacagctcgcacttttccggatggcttccaacctctggctctccgtc 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 tctccagctggttacagctcgcacttttccggatggcttccaacctctggctctccgtc 900

Qy    901 atcagctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaactatcgg 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 atcagctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaactatcgg 960

Qy    961 ttcctgcggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttat 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ttcctgcggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttat 1020

Qy   1021 atagttaatcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 atagttaatcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggt 1080

Qy   1081 tttattcggtgctcggtcaaagattggcacccatccgtctggatctatataagtggaccg 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 tttattcggtgctcggtcaaagattggcacccatccgtctggatctatataagtggaccg 1140

Qy   1141 tcgctggccctgagcactttcaatgtcgccatgttcgccctgacagccatttacattagg 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 tcgctggccctgagcactttcaatgtcgccatgttcgccctgacagccatttacattagg 1200

Qy   1201 aaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggattaactgcata 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 aaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggattaactgcata 1260

Qy   1261 aactttgacagccagacgtaagatatgataaaaccttggaccttggtaaactagttatcc 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 aactttgacagccagacgtaagatatgataaaaccttggaccttggtaaactagttatcc 1320

Qy   1321 ttgccgtattatccacatagttacctacagttcctgcggctctccatcgtgatgggcctt 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 ttgccgtattatccacatagttacctacagttcctgcggctctccatcgtgatgggcctt 1380

Qy   1381 acttggatattcaatgtcattccgtattctgcacggctccacattttctgggaatgggtc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 acttggatattcaatgtcattccgtattctgcacggctccacattttctgggaatgggtc 1440

Qy   1441 gggataatatccgagtattttcacagcgcgtttggaattgttttgttcgttctgctcgtc 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 gggataatatccgagtattttcacagcgcgtttggaattgttttgttcgttctgctcgtc 1500

Qy   1501 ctgaagcgcagcacatggactctaatgatggattct 1536
          |||||||||||||||||||||||||||||||||||
Db   1501 ctgaagcgcagcacatggactctaatgatggattct 1536
```

RE 8
U 191-681-13249
; Sequence 13249, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; TITLE OF INVENTION: USES THEREOF.
; FILE REFERENCE: c1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13249
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-60-191-681-13249

Query Match 100.0%; Score 1536; DB 51; Length 1536;
Local Similarity 100.0%; Pred. No. 0;
1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 atgaggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgaggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactca 60

Qy     61 aatgcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaagg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 aatgcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaagg 120

Qy    121 caaaacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactac 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 caaaacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactac 180

Qy    181 gagttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgt 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gagttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgt 240

Qy    241 gtttgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttggccaac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gtttgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttggccaac 300

Qy    301 ggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttc 360

Qy    361 acatacatggacctacaggcacgagtaccactcaccgatatggctattatcagagatgag 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 acatacatggacctacaggcacgagtaccactcaccgatatggctattatcagagatgag 420

Qy    421 ttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggtt 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggtt 480

Qy    481 agcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaattttgg 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 agcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaattttgg 540

Qy    541 gttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaactttgat 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 gttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaactttgat 600

Qy    601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660

Qy    661 ggatccttagagattctaattataacaatgatatgctttgtcctaacaatcgcagtatat 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ggatccttagagattctaattataacaatgatatgctttgtcctaacaatcgcagtatat 720

Qy    721 ctatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcg 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ctatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcg 780

Qy    781 aggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgc 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 aggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgc 840

Qy    841 tctccagctggttacagctcgcacttttccggatggcttccaacctctggctctccgtc 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 tctccagctggttacagctcgcacttttccggatggcttccaacctctggctctccgtc 900

Qy    901 atcagctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaactatcgg 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 atcagctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaactatcgg 960

Qy    961 ttcctgcggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttat 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ttcctgcggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttat 1020

Qy   1021 atagttaatcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 atagttaatcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggt 1080
```

This Page Blank (uspto)

ALIGNMENTS

```
RESULT   1
US-09-618-893-86
; Sequence 86, Application US/09618893
; GENERAL INFORMATION:
;  APPLICANT:  Cravchik, Anibal
;  TITLE OF INVENTION:  ISOLATED G-PROTEIN COUPLED RECEPTORS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF AS INSECTICIDAL TARGETS
;  FILE REFERENCE:  cl000733
;  CURRENT APPLICATION NUMBER: US/09/618,893
;  CURRENT FILING DATE:  2000-07-18
;  PRIOR APPLICATION NUMBER:  60/168,677
;  PRIOR FILING DATE:  1999-12-03
;  PRIOR APPLICATION NUMBER:  60/175,691
;  PRIOR FILING DATE:  2000-01-12
;  PRIOR APPLICATION NUMBER:  60/191,638
;  PRIOR FILING DATE:  2000-03-23
;  NUMBER OF SEQ ID NOS:  198
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 86
;  LENGTH: 1536
;  TYPE: DNA
;  ORGANISM: Drosophila
US-09-618-893-86

  Query Match             100.0%;  Score 1536;  DB 23;  Length 1536;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1536;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 atgaggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactca 60
Db       1 atgaggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactca 60

Qy      61 aatgcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaagg 120
Db      61 aatgcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaagg 120

Qy     121 caaaacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactac 180
Db     121 caaaacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactac 180

Qy     181 gagttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgt 240
Db      81 gagttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgt 240

Qy     241 gtttgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttggccaac 300
Db     241 gtttgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttggccaac 300

Qy     301 ggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttc 360
Db     301 ggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttc 360

Qy     361 acatacatggacctacaggcacgagtaccactcaccgatatggctattatcagagatgag 420
Db     361 acatacatggacctacaggcacgagtaccactcaccgatatggctattatcagagatgag 420

Qy     421 ttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggtt 480
Db     421 ttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggtt 480

Qy     481 agcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaatttttgg 540
Db     481 agcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaatttttgg 540

Qy     541 gttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaactttgat 600
Db     541 gttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaactttgat 600

Qy     601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660
Db     601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660

Qy     661 ggatccttagagattctaattataacaatgatatgctttgtcctaacaatcgcagtatat 720
Db     661 ggatccttagagattctaattataacaatgatatgctttgtcctaacaatcgcagtatat 720

Qy     721 ctatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcg 780
Db     721 ctatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcg 780

Qy     781 aggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgc 840
Db     781 aggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgc 840

Qy     841 tctccagctggttacagctcgcacttttccggatggcttccaacctctggctctccgtc 900
Db     841 tctccagctggttacagctcgcacttttccggatggcttccaacctctggctctccgtc 900

Qy     901 atcagctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaactatcgg 960
Db     901 atcagctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaactatcgg 960

Qy     961 ttcctgcggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttat 1020
Db     961 ttcctgcggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttat 1020

Qy    1021 atagttaatcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggt 1080
Db    1021 atagttaatcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggt 1080

Qy    1081 tttattcggtgctcggtcaaagattggcacccatccgtctggatctatataagtggaccg 1140
Db    1081 tttattcggtgctcggtcaaagattggcacccatccgtctggatctatataagtggaccg 1140

Qy    1141 tcgctggccctgagcactttcaatgtcgccatgttcgccctgacagccatttacattagg 1200
Db    1141 tcgctggccctgagcactttcaatgtcgccatgttcgccctgacagccatttacattagg 1200

Qy    1201 aaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggattaactgcata 1260
Db    1201 aaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggattaactgcata 1260

Qy    1261 aactttgacagccagacgtaagatatgataaaaccttggaccttggtaaactagttatcc 1320
Db    1261 aactttgacagccagacgtaagatatgataaaaccttggaccttggtaaactagttatcc 1320

Qy    1321 ttgccgtattatccacatagttacctacagttcctgcggctctccatcgtgatgggcctt 1380
Db    1321 ttgccgtattatccacatagttacctacagttcctgcggctctccatcgtgatgggcctt 1380

Qy    1381 acttggatattcaatgtcattccgtattctgcacggctccacattttctgggaatgggtc 1440
Db    1381 acttggatattcaatgtcattccgtattctgcacggctccacattttctgggaatgggtc 1440

Qy    1441 gggataatatccgagtattttcacagcgcgtttggaattgttttgttcgttctgctcgtc 1500
Db    1441 gggataatatccgagtattttcacagcgcgtttggaattgttttgttcgttctgctcgtc 1500

Qy    1501 ctgaagcgcagcacatggactctaatgatggattct 1536
Db    1501 ctgaagcgcagcacatggactctaatgatggattct 1536

RESULT   2
US-09-619-049-854
; Sequence 854, Application US/09619049
; GENERAL INFORMATION:
;  APPLICANT:  YANDELL, MARK
;  TITLE OF INVENTION:  ISOLATED DROSOPHILA PROTEINS ESSENTIAL
;  TITLE OF INVENTION:  FOR SURVIVAL, NUCLEIC ACID MOLECULES ENCODING ESSENTIAL
```

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 14:37:34 ; Search time 63.26 Seconds
(without alignments)
408.249 Million cell updates/sec

Title: US-09-619-049-855
Perfect score: 2324
Sequence: 1 MRLPWVIFCTVLLLIFTNNS.....NKFTNEEEGRINCINFDSQT 426

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601:*
1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	118	5.1	874	18	AAW07600	Human G-protein re
2	118	5.1	874	20	AAW94075	Human G-protein co
3	118	5.1	884	19	AAW59666	Amino acid sequenc
4	118	5.1	884	20	AAW94074	Human G-protein co
5	106	4.6	521	20	AAW30638	Partial human 7-tr
6	106	4.6	652	20	AAW30637	Human 7-transmembr
7	106	4.6	652	21	AAY59300	Human EGPCR polype
8	104.5	4.5	1013	21	AAB01247	Human HE6 receptor
9	104.5	4.5	1038	18	AAW36903	Human epididymis-s
10	102	4.4	417	20	AAW89226	Tumour necrosis fa
11	101.5	4.4	326	19	AAW44194	Arabidopsis G prot
12	101.5	4.4	326	21	AAG35727	Arabidopsis thalia
13	101.5	4.4	331	21	AAG05052	Arabidopsis thalia
14	101.5	4.4	350	21	AAG05051	Arabidopsis thalia
15	101.5	4.4	354	21	AAG35726	Arabidopsis thalia
16	100.5	4.3	420	20	AAW89224	Tumour necrosis fa
17	99.5	4.3	381	22	AAB75210	Drosophila gustato
18	99.5	4.3	449	22	AAB68879	Human RECAP polype
19	99	4.3	607	21	AAB38332	Human secreted pro
20	98.5	4.2	256	18	AAW33357	TBP(20-161)/hCG-al
21	98.5	4.2	540	21	AAY90903	Human frizzled-4 p
22	97.5	4.2	336	18	AAW33360	TBP(20-190)/hCG-be
23	97	4.2	611	20	AAW73924	Nucleobase permeas
24	96.5	4.2	453	21	AAY94340	Human cell surface
25	96.5	4.2	537	18	AAW31270	Mouse frizzled-4 p
26	96	4.1	886	22	AAB71869	Human EMR1 seven t
27	95.5	4.1	448	21	AAY51349	Sunflower HADES pr
28	95	4.1	1245	16	AAR70106	TNF-R-Pl. vivax Du
29	94.5	4.1	309	16	AAR70108	TNF-R-GBPH fusion
30	94.5	4.1	372	22	AAY72816	Calendula officina
31	94.5	4.1	547	16	AAR70104	TNF-R-GBPH fusion
32	94	4.0	1572	18	AAW27160	Mouse receptor ME2
33	94	4.0	2707	18	AAW27161	Mouse receptor ME2
34	92.5	4.0	264	21	AAG35728	Arabidopsis thalia
35	92.5	4.0	269	21	AAG05053	Arabidopsis thalia
36	92.5	4.0	446	20	AAW85122	A delta-6 desatura
37	92.5	4.0	448	17	AAR98455	Borage delta-6-des
38	92.5	4.0	448	19	AAW67471	Borage delta-6 des
39	92.5	4.0	448	20	AAW98130	Borage delta-6 des
40	91.5	3.9	681	21	AAY54070	Enzyme EPS6 which
41	91.5	3.9	681	21	AAY43772	Amino acid sequenc
42	91	3.9	307	18	AAW33358	TBP(20-161)/hCG-be
43	91	3.9	455	11	AAR07451	Human Tumour Necro
44	91	3.9	467	21	AAG36372	Arabidopsis thalia
45	91	3.9	592	21	AAG36371	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAW07600
ID AAW07600 standard; Protein; 874 AA.
XX
AC AAW07600;
XX
DT 26-FEB-1997 (first entry)
XX
DE Human G-protein receptor HCEGH45.
XX
KW G-protein receptor; HCEGH45; signal transduction;
KW pituitary adenylate cyclase activating polypeptide;
KW PACAP hypersecretion; Alzheimer's disease; Parkinson's disease;
KW amnesia; antagonist; agonist; diagnosis; therapy.
XX
OS Homo sapiens.
XX
PN WO9639439-A1.
XX
PD 12-DEC-1996.
XX
PF 06-JUN-1995; 95WO-US07188.
XX
PR 06-JUN-1995; 95WO-US07188.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Li Y, Rosen CA, Ruben SM, Soppet DR;
XX
DR WPI; 1997-043074/04.
XX
PT Human G-protein pituitary adenylate cyclase activating polypeptide
PT receptor, HCEGH45 - useful to identify (ant)agonists, for treatment
PT of asthma, amnesia, Parkinson's disease, etc.

```
XX
PS   Claim 1; Fig 1A-J; 65pp; English.
XX
CC   Novel human mature G-protein receptor HCEGH45 (AAW07600) is a
CC   G-protein pituitary adenylate cyclase activating polypeptide
CC   (PACAP)-like receptor of amnesiac-like neuropeptides. It is a
CC   7-transmembrane protein and participates in signal transduction
CC   pathways. Its amino acid sequence was deduced from a cDNA clone
CC   (AAT44039) derived from human cerebellum tissue. Recombinant HCEGH45
CC   can be produced in host (e.g. COS, Sf9) cells and used to identify
CC   (ant)agonists to the polypeptide. Antagonists can be used to treat
CC   PACAP hypersecretory conditions and to create amnesia models.
CC   Agonists are employed to treat amnesia and Alzheimer's disease.
XX
SQ   Sequence   874 AA;


  Query Match             5.1%;  Score 118;  DB 18;  Length 874;
  Best Local Similarity   22.1%;  Pred. No. 0.0036;
  Matches   50;  Conservative   33;  Mismatches   93;  Indels   50;  Gaps    9;

Qy     207 MWIIR-----HRCTSH----------------ISP-----GSLEIL--------IITMIC 232
           :|  |     ||  ||                ||     | ||:|         :::
Db      56 vwtardcelvhrngsharcrcsrtgtfgvlmdasprerlegdlellavfthvvvavsvaa 115

Qy     233 FVLTIAVYLYIKKLR-NVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFR 291
           ||| |: | :: |: || |           : | ::  |      :|:      |:|
Db     116 lvltaaillslrslksnvrgihanvaaalgvaellfllgihrthnqlvctavaillhyff 175

Qy     292 MASNLWLSVISYHTWKVLTSLNRVDPN-YRFLRYNAFVWSTAAIMTGSIYIVNQIWENDP 350
           :::  || |   | :::      ||    ||  |:|  |   |:: |    ::     :|
Db     176 lstfawlfvqglhlyrmqveprnvdrgamrf--yhalgwgvpavllglavgldpegygnp 233

Qy     351 SKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTA 396
                   | :    | :| : :|| : :   |  || | |
Db     234 d----------fcwisvhepliwsf-agpvvlvivmngtmfllaa 267


RESULT   2
AAW94075
ID   AAW94075 standard; Protein; 874 AA.
XX
AC   AAW94075;
XX
DT   19-APR-1999  (first entry)
XX
DE   Human G-protein coupled receptor polypeptide, HCEGH45.
XX
KW   Human; G-protein coupled receptor; HCEGH45; PACAP; neuropeptide;
KW   pituitary adenylate cyclase activating polypeptide; hypersecretory;
KW   amnesia; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 708
FT                   /label= unknown
FT                   /note= "encoded by NNN"
XX
PN   US5869632-A.
XX
PD   09-FEB-1999.
XX
PF   06-JUN-1995;   95US-0465976.
XX
PR   06-JUN-1995;   95US-0465976.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Li Y,  Rosen CA,  Ruben SM,  Soppet DR;
XX
DR   WPI; 1999-152883/13.
DR   N-PSDB; AAX01094.
XX
PT   DNA encoding human G-protein coupled receptor - useful for producing
PT   recombinant polypeptide
XX
PS   Disclosure; Fig 1A-L; 37pp; English.
XX
CC   This represents a human G-protein coupled receptor polypeptide, HCEGH45.
CC   The encoding cDNA is deposited under the accession no: ATCC 97132. The
CC   polynucleotide is used to produce a recombinant HCEGH45 which has been
CC   putatively identified as a human G-protein pituitary adenylate cyclase
CC   activating polypeptide (PACAP)-like receptor for amnesiac like
CC   neuropeptides. Antagonists of the polypeptide may be used therapeutically
CC   to treat PACAP hypersecretory conditions while agonists may be employed
CC   to treat amnesia and Alzheimer's disease.
CC   Note: this sequence (provided in the Figure 1A-L) differs from the
CC   HCEGH45 sequence given in the sequence listing after amino acid residue
CC   687.
XX
SQ   Sequence   874 AA;


  Query Match             5.1%;  Score 118;  DB 20;  Length 874;
  Best Local Similarity   22.1%;  Pred. No. 0.0036;
  Matches   50;  Conservative   33;  Mismatches   93;  Indels   50;  Gaps    9;

Qy     207 MWIIR-----HRCTSH----------------ISP-----GSLEIL--------IITMIC 232
           :|  |     ||  ||                ||     | ||:|         :::
Db      56 vwtardcelvhrngsharcrcsrtgtfgvlmdasprerlegdlellavfthvvvavsvaa 115

Qy     233 FVLTIAVYLYIKKLR-NVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFR 291
           ||| |: | :: |: || |           : | ::  |      :|:      |:|
Db     116 lvltaaillslrslksnvrgihanvaaalgvaellfllgihrthnqlvctavaillhyff 175

Qy     292 MASNLWLSVISYHTWKVLTSLNRVDPN-YRFLRYNAFVWSTAAIMTGSIYIVNQIWENDP 350
           :::  || |   | :::      ||    ||  |:|  |   |:: |    ::     :|
Db     176 lstfawlfvqglhlyrmqveprnvdrgamrf--yhalgwgvpavllglavgldpegygnp 233

Qy     351 SKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTA 396
                   | :    | :| : :|| : :   |  || | |
Db     234 d----------fcwisvhepliwsf-agpvvlvivmngtmfllaa 267


RESULT   3
AAW59666
ID   AAW59666 standard; Protein; 884 AA.
XX
AC   AAW59666;
XX
DT   12-OCT-1998  (first entry)
XX
DE   Amino acid sequence of human G-protein coupled receptor.
XX
KW   Human; G-protein coupled receptor; HCEGH45; hypersecretory condition;
KW   pituitary adenylate cyclase; agonist; memory; amnesia;
KW   nerve cell death; neuropathy; Alzheimer's disease; antagonist.
XX
OS   Homo sapiens.
XX
PN   WO9824900-A1.
XX
PD   11-JUN-1998.
XX
PF   21-NOV-1997;   97WO-US20547.
XX
PR   02-DEC-1996;   96US-0032186.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Li Y,  Rosen CA,  Ruben SM,  Soppet DR;
XX
DR   WPI; 1998-333320/29.
DR   N-PSDB; AAV41550.
XX
```

PT New isolated human G-protein receptor, HCEGH45 - used to develop
PT products for treating e.g. hypersecretory conditions, to improve
PT memory, to treat amnesia or to prevent nerve cell death in
PT neuropathy
XX
PS Claim 4; Fig 1A-1D; 82pp; English.
XX
CC This is the amino acid sequence of the human G-protein receptor
CC (HCEGH45) used in the method of the invention to develop products to
CC treat hypersecretory conditions. The novel protein is a human
CC G-protein pituitary adenylate cyclase activating polypeptide
CC (PACAP)-like receptor for amneslac like neuropeptides, designated
CC HCEGH45. Agonists can be used to treat hyposecretory conditions, to
CC improve memory, to treat amnesia and prevent nerve cell death in
CC neuropathy to prevent and/or treat diseases such as Alzheimer's
CC disease. G-protein coupled receptor agonists can also be used to
CC treat asthma, Parkinson's disease, acute heart failure, hypotension,
CC urinary retention, and osteoporosis. Antagonists can be used to treat
CC hypersecretory conditions and to create pharmacological amnesia or
CC effect long-term memory. G-protein coupled receptor antagonists can
CC also be used to treat hypertension, angina pectoris, myocardial
CC infarction, ulcers, asthma, allergies, psychoses, depression, migraine,
CC vomiting, and benign prostatic hypertrophy. The products can also be
CC used for detection, diagnosis and drug screening.
XX
SQ Sequence 884 AA;

Query Match 5.1%; Score 118; DB 19; Length 884;
Best Local Similarity 22.1%; Pred. No. 0.0037;
Matches 50; Conservative 33; Mismatches 93; Indels 50; Gaps 9;

```
Qy     207 MWIIR-----HRCTSH---------------ISP-----GSLEIL--------IITMIC 232
           :|  |     ||  ||                 ||     | ||:|         :::
Db      56 vwtardcelvhrngsharcrcsrtgtfgvlmdasprerlegdlellavfthvvvavsvaa 115

Qy     233 FVLTIAVYLYIKKLR-NVTGKCIVCCIVSRFIQCLIMILDHLNLLLNGICSPAGYSSHFFR 291
           |||  |: | :: |: || |           : | ::  |      :|:      |:|
Db     116 lvltaaillslrslksnvrgihanvaaalgvaellfllgihrthnqlvctavaillhyff 175

Qy     292 MASNLWLSVISYHTWKVLTSLNRVDPN-YRFLRYNAFVWSTAAIMTGSIYIVNQIWENDP 350
           :::  || |   | :::      ||    ||  |:|  |   |:: |    ::     :|
Db     176 lstfawlfvqglhlyrmqveprnvdrgamrf--yhalgwgvpavllglavgldpegygnp 233

Qy     351 SKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTA 396
                 | :    | :| : :|| : :   |  || | |
Db     234 d----------fcwisvhepliwsf-agpvvlvivmngtmfllaa 267
```

RESULT 4
AAW94074
ID AAW94074 standard; Protein; 884 AA.
XX
AC AAW94074;
XX
DT 19-APR-1999 (first entry)
XX
DE Human G-protein coupled receptor polypeptide, HCEGH45.
XX
KW Human; G-protein coupled receptor; HCEGH45; PACAP; neuropeptide;
KW pituitary adenylate cyclase activating polypeptide; hypersecretory;
KW amnesia; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN US5869632-A.
XX
PD 09-FEB-1999.
XX
PF 06-JUN-1995; 95US-0465976.
XX
PR 06-JUN-1995; 95US-0465976.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Li Y, Rosen CA, Ruben SM, Soppet DR;
XX
DR WPI; 1999-152883/13.
DR N-PSDB; AAX01089.
XX
PT DNA encoding human G-protein coupled receptor - useful for producing
PT recombinant polypeptide
XX
PS Claim 1; Columns 21-30; 37pp; English.
XX
CC This represents a human G-protein coupled receptor polypeptide, HCEGH45.
CC The encoding cDNA is deposited under the accession no: ATCC 97132. The
CC polynucleotide is used to produce a recombinant HCEGH45 which has been
CC putatively identified as a human G-protein pituitary adenylate cyclase
CC activating polypeptide (PACAP)-like receptor for amnesiac like
CC neuropeptides. Antagonists of the polypeptide may be used therapeutically
CC to treat PACAP hypersecretory conditions while agonists may be employed
CC to treat amnesia and Alzheimer's disease.
CC Note: this sequence (provided in the sequence listing) differs from the
CC HCEGH45 sequence given in Fig 1A-L after amino acid residue 687.
XX
SQ Sequence 884 AA;

Query Match 5.1%; Score 118; DB 20; Length 884;
Best Local Similarity 22.1%; Pred. No. 0.0037;
Matches 50; Conservative 33; Mismatches 93; Indels 50; Gaps 9;

```
Qy     207 MWIIR-----HRCTSH---------------ISP-----GSLEIL--------IITMIC 232
           :|  |     ||  ||                 ||     | ||:|         :::
Db      56 vwtardcelvhrngsharcrcsrtgtfgvlmdasprerlegdlellavfthvvvavsvaa 115

Qy     233 FVLTIAVYLYIKKLR-NVTGKCIVCCIVSRFIQCLIMILDHLNLLLNGICSPAGYSSHFFR 291
           |||  |: | :: |: || |           : | ::  |      :|:      |:|
Db     116 lvltaaillslrslksnvrgihanvaaalgvaellfllgihrthnqlvctavaillhyff 175

Qy     292 MASNLWLSVISYHTWKVLTSLNRVDPN-YRFLRYNAFVWSTAAIMTGSIYIVNQIWENDP 350
           :::  || |   | :::      ||    ||  |:|  |   |:: |    ::     :|
Db     176 lstfawlfvqglhlyrmqveprnvdrgamrf--yhalgwgvpavllglavgldpegygnp 233

Qy     351 SKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTA 396
                 | :    | :| : :|| : :   |  || | |
Db     234 d----------fcwisvhepliwsf-agpvvlvivmngtmfllaa 267
```

RESULT 5
AAW30638
ID AAW30638 standard; Protein; 521 AA.
XX
AC AAW30638;
XX
DT 01-APR-1999 (first entry)
XX
DE Partial human 7-transmembrane receptor HAPO167 protein.
XX
KW Human; HAPO167; 7-transmembrane receptor; infection; bacterial; viral;
KW fungal; protozoan; HIV-1; HIV-2; pain; cancer; bulimia; asthma; anorexia;
KW G-protein coupled receptor; Parkinson's disease; acute heart failure;
KW hypotension; hypertension; urinary retention; osteoporosis; ulcer;
KW allergy; angina pectoris; myocardial infarction; psychotic disorder;
KW benign prostatic hypertrophy; neurological disorder.
XX
OS Homo sapiens.
XX
PN EP887407-A2.
XX
PD 30-DEC-1998.
XX
PF 05-MAY-1998; 98EP-0303521.
XX
PR 23-OCT-1997; 97US-0956322.

PR 11-JUN-1997; 97US-0049329.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Mao JY, Sathe GM;
XX
DR WPI; 1999-047874/05.
DR N-PSDB; AAX03845.
XX
PT New G-protein coupled receptor HAPO167 polypeptide and
PT polynucleotide - useful as diagnostic reagents and for prevention
PT and treatment of cancer, HIV infections and osteoporosis
XX
PS Disclosure; Page 13; 30pp; English.
XX
CC The present sequence is a partial G-protein coupled receptor (HAPO167)
CC protein. HAPO167 proteins and polynucleotides are useful for diagnosing
CC susceptibility to diseases by detecting mutations in the HAPO167 gene,
CC and can diagnose diseases associated with HAPO167 imbalance by
CC determining HAPO167 protein levels. HAPO167 proteins can be used to
CC screen for agonists and antagonists which bind to the protein. These can
CC be used in treatment to activate (agonist) or inhibit (antagonist)
CC HAPO167 activity, in addition to direct administration of antisense
CC sequences to prevent expression, or HAPO167 polynucleotides to treat
CC conditions associated with a lack of HAPO167 protein. Gene therapy may
CC also be used to affect endogenous HAPO167 protein production. HAPO167
CC antibodies are useful for inducing an immune response to immunise and
CC protect against diseases, and for isolating HAPO167 clones or purifying
CC the protein by affinity chromatography. HAPO167 proteins can be
CC administered directly or as a vaccine to inoculate against disease.
CC Diseases diagnosed, prevented or treated include: infections caused by
CC HIV-1 or -2; pain; cancer; bulimia; asthma; Parkinson's disease; acute
CC heart failure; hypotension; hypertension; urinary retention;
CC osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies;
CC benign prostatic hypertrophy; and psychotic and neurological disorders,
CC including delirium, anxiety, schizophrenia, manic depression, dementia,
CC severe mental retardation and dyskinesias, such as Huntington's disease
CC or Gilles de la Tourette's syndrome.
XX
SQ Sequence 521 AA;

Query Match 4.6%; Score 106; DB 20; Length 521;
Best Local Similarity 20.0%; Pred. No. 0.03;
Matches 54; Conservative 41; Mismatches 93; Indels 82; Gaps 13;

```
Qy     173 WFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEI-----LI 227
           | :|| | : |:   :   |   |       | : :   ||      | :     |
Db     185 wsrdgcflihvn---kshtmcncsh-------lssfavlmaltsqeedpvltvityvgls 234

Qy     228 ITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQ-CLIMILDHLNLLNGI------- 279
           ::::| :|    :|  | ::| :         |  :|  | : | ||  | ||
Db     235 vsllclllaaltfllckaiqnts---------tslhlqlslclflahllflvgidrtepkv 286

Qy     280 -CSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGS 338
            ||    : |:  :|:  |: :   | :  ||: |    ||  :
Db     287 lcsiiagalhylylaaftwmllegvhlf--ltarnltvvnyssi---------------- 328

Qy     339 IYIVNQIWENDPSKWNWLPLVGF------IRCSVKDWHPSV-------WIYIS------- 378
                 |   ||   | ||:      :  |   | | :       |:::
Db     329 ---------nrlmkwimfp-vgygvpavtvaisaasw-phlygtadrcwlhldqgfmwsf 377

Qy     379 -GPSLALSTFNVAMFALTAIYIRKVKGGIN 407
            ||  |: : |: :| |    :::    :|
Db     378 lgpvcaifsanlvlfilvfwilkrklssln 407
```

RESULT 6
AAW30637
ID AAW30637 standard; Protein; 652 AA.
XX
AC AAW30637;
XX

DT 01-APR-1999 (first entry)
XX
DE Human 7-transmembrane receptor HAPO167 protein.
XX
KW Human; HAPO167; 7-transmembrane receptor; infection; bacterial; viral;
KW fungal; protozoan; HIV-1; HIV-2; pain; cancer; bulimia; asthma; anorexia;
KW G-protein coupled receptor; Parkinson's disease; acute heart failure;
KW hypotension; hypertension; urinary retention; osteoporosis; ulcer;
KW allergy; angina pectoris; myocardial infarction; psychotic disorder;
KW benign prostatic hypertrophy; neurological disorder.
XX
OS Homo sapiens.
XX
PN EP887407-A2.
XX
PD 30-DEC-1998.
XX
PF 05-MAY-1998; 98EP-0303521.
XX
PR 23-OCT-1997; 97US-0956322.
PR 11-JUN-1997; 97US-0049329.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Mao JY, Sathe GM;
XX
DR WPI; 1999-047874/05.
DR N-PSDB; AAX03844.
XX
PT New G-protein coupled receptor HAPO167 polypeptide and
PT polynucleotide - useful as diagnostic reagents and for prevention
PT and treatment of cancer, HIV infections and osteoporosis
XX
PS Claim 11; Page 23-25; 30pp; English.
XX
CC The present sequence represents a G-protein coupled receptor (HAPO167)
CC protein. HAPO167 proteins and polynucleotides are useful for diagnosing
CC susceptibility to diseases by detecting mutations in the HAPO167 gene,
CC and can diagnose diseases associated with HAPO167 imbalance by
CC determining HAPO167 protein levels. HAPO167 proteins can be used to
CC screen for agonists and antagonists which bind to the protein. These can
CC be used in treatment to activate (agonist) or inhibit (antagonist)
CC HAPO167 activity, in addition to direct administration of antisense
CC sequences to prevent expression, or HAPO167 polynucleotides to treat
CC conditions associated with a lack of HAPO167 protein. Gene therapy may
CC also be used to affect endogenous HAPO167 protein production. HAPO167
CC antibodies are useful for inducing an immune response to immunise and
CC protect against diseases, and for isolating HAPO167 clones or purifying
CC the protein by affinity chromatography. HAPO167 proteins can be
CC administered directly or as a vaccine to inoculate against disease.
CC Diseases diagnosed, prevented or treated include: infections caused by
CC HIV-1 or -2; pain; cancer; bulimia; asthma; Parkinson's disease; acute
CC heart failure; hypotension; hypertension; urinary retention;
CC osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies;
CC benign prostatic hypertrophy; and psychotic and neurological disorders,
CC including delirium, anxiety, schizophrenia, manic depression, dementia,
CC severe mental retardation and dyskinesias, such as Huntington's disease
CC or Gilles de la Tourette's syndrome.
XX
SQ Sequence 652 AA;

Query Match 4.6%; Score 106; DB 20; Length 652;
Best Local Similarity 20.0%; Pred. No. 0.041;
Matches 54; Conservative 41; Mismatches 93; Indels 82; Gaps 13;

```
Qy     173 WFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEI-----LI 227
           | :|| | : |:   :   |   |       | : :   ||      | :     |
Db     316 wsrdgcflihvn---kshtmcncsh-------lssfavlmaltsqeedpvltvityvgls 365

Qy     228 ITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQ-CLIMILDHLNLLNGI------- 279
           ::::| :|    :|  | ::| :         |  :|  | : | ||  | ||
Db     366 vsllclllaaltfllckaiqnts--------tslhlqlslclflahllflvgidrtepkv 417
```

```
Qy     280 -CSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGS 338
           ||    : |:  :|:  |: :   | :  ||: |    ||  :
Db     418 lcsiiagalhylylaaftwmllegvhlf--ltarnltvvnyssi---------------- 459

Qy     339 IYIVNQIWENDPSKWNWLPLVGF------IRCSVKDWHPSV-------WIYIS------- 378
                 |   ||   | ||:      :  |   | | :       |:::
Db     460 ---------nrlmkwimfp-vgygvpavtvaisaasw-phlygtadrcwlhldqgfmwsf 508

Qy     379 -GPSLALSTFNVAMFALTAIYIRKVKGGIN 407
            ||  |: : |: :| |    :::    :|
Db     509 lgpvcaifsanlvlfilvfwilkrklssln 538
```

RESULT 7
AAY59300
ID AAY59300 standard; Protein; 652 AA.
XX
AC AAY59300;
XX
DT 03-MAY-2000 (first entry)
XX
DE Human EGPCR polypeptide.
XX
KW Emr1-like G protein coupled receptor; EGPCR; gene therapy; cancer;
KW respiratory; antiinflammatory; immunomodulatory; cytostatic; human;
KW AIDS; antiasthmatic; antiviral; antibacterial; antifungal; antiparasitic;
KW antiarthritic; antiatherosclerotic; antianemic; antidiabetic;
KW dermalogical; antisclerotic; cardioactive.
XX
OS Homo sapiens.
XX
PN US6013479-A.
XX
PD 11-JAN-2000.
XX
PF 02-JUL-1998; 98US-0110116.
XX
PR 02-JUL-1998; 98US-0110116.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Xu H, Cohan VL, Stuart SG;
XX
DR WPI; 2000-146549/13.
DR N-PSDB; AAZ40700.
XX
PT Polynucleotide encoding human Emr1-like G protein coupled receptor
PT polypeptide is useful for producing recombinant polypeptide or as a
PT probe -
XX
PS Claim 1; Fig 1A-H; 43pp; English.
XX
CC This represents a human Emr1-like G protein coupled receptor polypeptide
CC (EGPCR). Recombinant EGPCR can be expressed by standard recombinant
CC methodology and can be used to treat or prevent a disorder associated
CC with decreased expression or activity of EGPCR, e.g. a respiratory
CC disorder (e.g. allergies, asthma, acute inflammatory lung disease,
CC chronic inflammatory lung disease, chronic obstructive pulmonary
CC dysplasia, emphysema, adult respiratory distress syndrome, bronchitis
CC and intestinal lung diseases), complications of cancer, hemodialysis and
CC extracorporeal circulation, bacterial, fungal, parasitic, protozoal
CC viral and helminthic infections and trauma, an inflammatory disorder
CC (e.g. rheumatoid arthritis, osteoarthritis, inflammatory bowel disease,
CC adult respiratory distress syndrome, allergies, asthma, bronchitis,
CC cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis,
CC emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis
CC fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis,
CC Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis,
CC hypereosinophilia, intestitial lung disease, myasthenia gravis,
CC myocardial or pericardial inflammation, pancreatitis, polymyositis,
CC psoriasis, Reiter's syndrome, scleroderma, Sjogren's syndrome, synstemic
CC anaphylaxis, systemic sclerosis, thrombocytopenic purpura, ulcerative
CC colitis, uveitis, Werner syndrome) and an immunological disorder (e.g.
CC acquired immunodeficiency syndrome (AIDS), Addison's disease, allergies,
CC ankylosing spondylitis, amyloidosis, anemia, atherosclerosis, autoimmune
CC hemolytic anemia, autoimmune throiditis, bronchitis and osteoporosis. The
CC polynucleotide can also be used as a probe for detecting other Emr-1 like
CC G protein coupled receptor genes.
XX
SQ Sequence 652 AA;

Query Match 4.6%; Score 106; DB 21; Length 652;
Best Local Similarity 20.0%; Pred. No. 0.041;
Matches 54; Conservative 41; Mismatches 93; Indels 82; Gaps 13;

```
Qy     173 WFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEI-----LI 227
           | :|| | : |:   :    |   |         | : :    ||     | :     |
Db     316 wsrdgcflihvn---kshtmcncsh-------lssfavlmaltsqeedpvltvityvgls 365

Qy     228 ITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQ-CLIMILDHLNLLNGI------- 279
           ::::| :|    :|  | ::| :          |  :|  | : | ||  | ||
Db     366 vsllclllaaltfllckaiqnts--------tslhlqlslclflahllflvgidrtepkv 417

Qy     280 -CSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGS 338
            ||    : |:  :|:  |: :   | :  ||: |    ||  :
Db     418 lcsiiagalhylylaaftwmllegvhlf--ltarnltvvnyssi---------------- 459

Qy     339 IYIVNQIWENDPSKWNWLPLVGF------IRCSVKDWHPSV-------WIYIS------- 378
                 |   ||   | ||:      :  |   | | :       |:::
Db     460 ---------nrlmkwimfp-vgygvpavtvaisaasw-phlygtadrcwlhldqgfmwsf 508

Qy     379 -GPSLALSTFNVAMFALTAIYIRKVKGGIN 407
            ||  |: : |: :| |    :::    :|
Db     509 lgpvcaifsanlvlfilvfwilkrklssln 538
```

RESULT 8
AAB01247
ID AAB01247 standard; Protein; 1013 AA.
XX
AC AAB01247;
XX
DT 03-OCT-2000 (first entry)
XX
DE Human HE6 receptor.
XX
KW Human; HE6 receptor; zsig56; seven transmembrane domain receptor;
KW hypotensive; antigout; cytostatic; antiinflammatory; ion homeostasis;
KW cardiant; neurotransmitter; neuroprotective; antiparkinsonian;
KW cerebroprotective; nootropic; neuroleptic; tranquiliser; antiarthritic;
KW reproductive; signal transduction activator; bone disease; hypertension;
KW renal failure; heart failure; hyperthyroidism; hyperparathyroidism;
KW carcinoma; sarcoidosis; pancreatitis; stress; high blood pressure;
KW immune depression; periodontal disease; neurodegenerative disease;
KW multiple sclerosis, Alzheimer's disease; Parkinson's disease;
KW schizophrenia; manic depression; stroke; rheumatoid arthritis;
KW male fertility; spermatogenesis stimulation; pregnancy regulation;
KW prostate cancer.
XX
OS Homo sapiens.
XX
PN WO200034473-A2.
XX
PD 15-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28492.
XX
PR 10-DEC-1998; 98US-0208691.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sheppard PO, Ellsworth JL;
XX
DR WPI; 2000-442164/38.

```
XX
PT   Novel G-protein coupled receptor zsig56 useful for treating
PT   hypertension, hyper and hypothyroidism, inflammation, gout, carcinoma,
PT   pancreatitis, Alzheimer's disease and Parkinson's disease, renal and
PT   heart failure -
XX
PS   Disclosure; Fig 1; 121pp; English.
XX
CC   The present sequence is the human HE6 receptor, which shows homology to
CC   a seven transmembrane domain receptor designated zsig56. The full
CC   length nucleotide sequence was obtained from a human retina library.
CC   zsig56 polypeptides, nucleic acid, agonists and/or antagonists may be
CC   used to treat a wide range of disorders including certain bone diseases,
CC   hypertension, renal failure, gout, congestive heart failure,
CC   hyperthyroidism, hyperparathyroidism, certain carcinomas, sarcoidosis
CC   and pancreatitis. They can be used to treat disorders associated with
CC   changes in ion or electrolyte homeostasis, and stress induced disorders
CC   such as high blood pressure, heart failure, immune depression and
CC   periodontal disease. They may be used to treat neurodegenerative
CC   diseases, including multiple sclerosis, Alzheimer's disease and
CC   Parkinson's disease, schizophrenia and manic depression, and to repair
CC   nerve tissue following damage due to strokes and brain and spinal
CC   injuries. Inflammatory disorders such as rheumatoid arthritis can also
CC   be treated. zsig56 is expressed in tissues associated with reproduction,
CC   i.e. the testis, prostate and placenta, and may be used to treat male
CC   fertility by stimulating spermatogenesis and to regulate gestation and
CC   birth. zsig56 may be useful as a marker or therapeutic agent in the
CC   treatment of prostate cancer.
XX
SQ   Sequence   1013 AA;


  Query Match             4.5%;  Score 104.5;  DB 21;  Length 1013;
  Best Local Similarity   15.7%;  Pred. No. 0.11;
  Matches   32;  Conservative   48;  Mismatches   89;  Indels   35;  Gaps    5;

Qy     224 EILIITMICFV----------LTIAVYLYIKKL-RNVTGKCIVCCIVSRFIQCLIMILDH 272
           ::: :| | ::          :|:  |:  :|: |:   | ::    :  :  |: :||
Db     620 qmmaltfityigcglssiflsvtlvtyiafekirrdypskiliqlcaalllnlvfllds 679

Qy     273 ---LNLLNGICSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNAFVW 329
               | : |:|       |:| : |  |: : ::| :  |  :         |::    |
Db     680 wialykmqglcisvavflhyfllvsftwmgleafhmylalvkvfntyirkyilkfcivgw 739

Qy     330 STAAIMTGSIYIVN----------QIWENDPSKWNWLPLVGFIRCSVKDWHPSVWIYISG 379
             |::   |  ::          :     |  : |:          : :   :| : |
Db     740 gvpavvvtiiltispdnyglgsygkfpngspddfcwi----------nnnavfyitvvg 788

Qy     380 PSLALSTFNVAMFALTAIYIRKVK 403
               :   ||:|| :  : : ::|
Db     789 yfcvifllnvsmfivvlvqlcrik 812


RESULT   9
AAW36903
ID   AAW36903 standard; Protein; 1038 AA.
XX
AC   AAW36903;
XX
DT   11-MAY-1998  (first entry)
XX
DE   Human epididymis-specific receptor protein.
XX
KW   Epididymis-specific receptor protein; transmembrane protein;
KW   sperm maturation; male infertility; agonist; antagonist;
KW   contraception; autoantibody; diagnostic; detection; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..1038
FT                   /note= "partial protein sequence"
XX
PN   DE19617940-A1.
XX
PD   30-OCT-1997.
XX
PF   29-APR-1996;   96DE-1017940.
XX
PR   29-APR-1996;   96DE-1017940.
XX
PA   (IHFH-) IHF INST HORMON & FORTPFLANZUNGS.
PA   (HORM-) INST HORMON & FORTPFLANZUNGSFORSCHUNG GM.
XX
PI   Ivell R,  Osterhoff C;
XX
DR   WPI; 1997-527841/49.
DR   N-PSDB; AAT97955.
XX
PT   DNA encoding epididymis-specific receptor protein - useful for
PT   developing male infertility or contraceptive treatments
XX
PS   Claim 1; Page 19-22; 33pp; German.
XX
CC   This sequence represents a novel epididymis-specific receptor protein
CC   which is a transmembrane protein associated with the epididymis
CC   epithelium and is closely connected with sperm maturation. The protein
CC   or its derivative or fragments may be useful for detecting autoantibodies
CC   in the serum of infertile men and for developing ligands specific for the
CC   receptor, e.g. agonists that may stimulate sperm maturation and thus be
CC   useful for treating male infertility or antagonists that may be useful
CC   for contraception. The antibodies can be used to detect the receptor in
CC   vitro or in vivo. Oligonucleotides derived from the DNA sequences can
CC   be used as diagnostic probes.
XX
SQ   Sequence   1038 AA;


  Query Match             4.5%;  Score 104.5;  DB 18;  Length 1038;
  Best Local Similarity   15.7%;  Pred. No. 0.11;
  Matches   32;  Conservative   48;  Mismatches   89;  Indels   35;  Gaps    5;

Qy     224 EILIITMICFV----------LTIAVYLYIKKL-RNVTGKCIVCCIVSRFIQCLIMILDH 272
           ::: :| | ::          :|:  |:  :|: |:   | ::    :  :  |: :||
Db     645 qmmaltfityigcglssiflsvtlvtyiafekirrdypskiliqlcaalllnlvfllds 704

Qy     273 ---LNLLNGICSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNAFVW 329
               | : |:|       |:| : |  |: : ::| :  |  :         |::    |
Db     705 wialykmqglcisvavflhyfllvsftwmgleafhmylalvkvfntyirkyilkfcivgw 764

Qy     330 STAAIMTGSIYIVN----------QIWENDPSKWNWLPLVGFIRCSVKDWHPSVWIYISG 379
             |::   |  ::          :     |  : |:          : :   :| : |
Db     765 gvpavvvtiiltispdnyglgsygkfpngspddfcwi----------nnnavfyitvvg 813

Qy     380 PSLALSTFNVAMFALTAIYIRKVK 403
               :   ||:|| :  : : ::|
Db     814 yfcvifllnvsmfivvlvqlcrik 837


RESULT  10
AAW89226
ID   AAW89226 standard; Protein; 417 AA.
XX
AC   AAW89226;
XX
DT   04-MAR-1999  (first entry)
XX
DE   Tumour necrosis factor bp/osteoprotegerin construct TNFbp/196.
XX
KW   Tumour necrosis factor receptor 1; TNFR-1; inhibitor; osteoprotegerin;
KW   OPG; chimeric; fusion; dimerisation domain; autoimmune disease;
KW   inflammation; apoptosis.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
```

```
PN   WO9849305-A1.
XX
PD   05-NOV-1998.
XX
PF   29-APR-1998;   98WO-US08631.
XX
PR   01-MAY-1997;   97US-0850188.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Boyle WJ,  Wooden S;
XX
DR   WPI; 1999-034661/03.
XX
PT   New chimeric osteoprotegerin polypeptides - contain the
PT   osteoprotegerin dimerisation domain and a heterologous sequence,
PT   useful to treat TNF and TNFR-mediated disorders
XX
PS   Example 1; Fig 4; 92pp; English.
XX
CC   The present invention describes a chimeric polypeptide (A1), comprising
CC   an osteoprotegerin (OPG) dimerisation domain fused to a heterologous
CC   amino acid sequence. Also described are: (1) a multimer polypeptide
CC   comprising covalently associated A1 monomers; (2) an isolated nucleic
CC   acid encoding A1; (3) an expression vector comprising the nucleic acid
CC   sequence; and (4) a host cell transformed or transfected with the
CC   expression vector so that the nucleic acid is expressible. The products
CC   from the present invention are useful to treat a variety of disorders
CC   including those related to receptor binding. Compositions comprising
CC   tumour necrosis factor (TNF)/OPG and TNF receptor (TNFR)/OPG chimeras
CC   are used to treat TNF and TNFR-mediated disorders such as inflammation,
CC   autoimmune diseases and disorders related to excessive apoptosis. The
CC   chimeras are also useful for detecting molecules which interact with
CC   fused heterologous sequences to identify potential new receptors and
CC   ligands. The present sequence represents a TNFbp/OPG construct from
CC   the example of the present invention for creating TNFbp/OPG fusion
CC   proteins.
XX
SQ   Sequence   417 AA;

  Query Match             4.4%;  Score 102;  DB 20;  Length 417;
  Best Local Similarity  20.4%;  Pred. No. 0.056;
  Matches   60;  Conservative   35;  Mismatches   67;  Indels  132;  Gaps   17;

Qy      37 YIERQNDS----------YLYDDIEIPASLTGYYEFRQFGDGSITPIEKHLRACV-CSVR 85
           ||  ||:|          |||:|   |   |   : |:   || |  | ||| |: ||
Db      49 yihpqnnsicctkchkgtylyndcpgpgqdt---dcrecesgsftasenhlrhciscs-- 103

Qy      86 PCIRICCPAKNFLANGKCDDGLKEELARFKPYIYFTYMDLQARVPLTDMAIIRDEFFDCD 145
                      ||    ::|:               :| ::   : ||    |
Db     104 ----------------kc----rkem--------------gqveissctvdrdtvcgcr 128

Qy     146 EMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFWVTVDL-FMEKQDYCLYRHNFDSDFP 204
           :    : : ::  |   | || | | :   :|    || |   |||:
Db     129 k----nqyrhywsenlfqcfn-csicl---ng----tvhlscqekqntv----------- 165

Qy     205 KSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQ 264
                  || |                          ::::      :|: |    : ::
Db     166 ---------ctchag--------------------fflre-----necvscsnckksle 190

Qy     265 CLIMILDHLNLLNG-----------ICSPAGYSSHFFRMA-----SNLWLSVI 301
           |  : |  :  : |          :|  |     ||| |     :  ||||:
Db     191 ctklclpqienvkgtedsgttgidvtlceea-----ffrfavptkftpnwlsvl 239


RESULT  11
AAW44194
ID   AAW44194 standard; Protein; 326 AA.
XX
AC   AAW44194;
XX
DT   06-JUL-1998  (first entry)
XX
DE   Arabidopsis G protein-coupled receptor AtGCR1.
XX
KW   G protein-coupled receptor; AtGCR1 gene; signal transduction;
KW   transgenic plant; cytokinin.
XX
OS   Arabidopsis thaliana var. Colombia.
XX
FH   Key             Location/Qualifiers
FT   Domain          25..44
FT                   /note= "transmembrane domain"
FT   Domain          54..72
FT                   /note= "transmembrane domain"
FT   Domain          85..105
FT                   /note= "transmembrane domain"
FT   Domain          121..142
FT                   /note= "transmembrane domain"
FT   Domain          161..185
FT                   /note= "transmembrane domain"
FT   Domain          218..238
FT                   /note= "transmembrane domain"
FT   Domain          248..272
FT                   /note= "transmembrane domain"
XX
PN   WO9800545-A2.
XX
PD   08-JAN-1998.
XX
PF   30-JUN-1997;   97WO-GB01766.
XX
PR   25-MAR-1997;   97GB-0006118.
PR   28-JUN-1996;   96GB-0013617.
XX
PA   (ARAB-) INST ARABLE CROPS RES.
XX
PI   Hooley R,  Plakidou-Dymock S;
XX
DR   WPI; 1998-086970/08.
DR   N-PSDB; AAV12369-70.
XX
PT   Recombinant or isolated DNA encoding protein having homology to G
PT   protein-coupled receptor in plants - used to develop products to
PT   influence the sensitivity of plants to cytokinins or other
PT   signalling compounds
XX
PS   Claim 6; Fig 2; 62pp; English.
XX
CC   This Arabidopsis thaliana protein, designated AtGCR1, is a plant
CC   homologue of G protein-coupled receptors (GPCR) that influences
CC   cytokinin sensitivity in roots and shoots.  Its amino acid sequence
CC   was deduced from genomic and cDNA clones (see AAV12369-70) obtained
CC   by homology to known GPCR sequences.  Genetic constructs containing
CC   AtGCR1 DNA in sense or antisense orientation can be used to alter G
CC   protein signal transduction in plants, thereby influencing plant
CC   growth, development and responses to endogenous signalling
CC   molecules, environmental signals, pests and pathogens.  The AtGCR1
CC   protein can be expressed in host cells and used to screen for
CC   potential agonists and antagonists.
XX
SQ   Sequence   326 AA;

  Query Match             4.4%;  Score 101.5;  DB 19;  Length 326;
  Best Local Similarity  22.1%;  Pred. No. 0.044;
  Matches   31;  Conservative   30;  Mismatches   68;  Indels   11;  Gaps    5;

Qy     218 ISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLN 277
           |: |:  :  :     ||     |   |:||  : | :     :|  :     :|:   :
Db      22 intgasslsfvgsafivl---cyclfkelrkfsfklvfylalsdmlcsfflivgdps-kg 77

Qy     278 GICSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTG 337
                ||   ||::||| :|| || : |::   : :   :: |       :: :|| |: ::|
Db      78 ficyaqgytthffcvasflwtttiaftlhrtvvk-hktdvedleamfhlyvwgtslvvt- 135
```

```
Qy    338 SIYIVNQIWENDP--SKWNW 355
              ::      |      | |
Db    136 ---virsfgnnhshlgpwcw 152


RESULT  12
AAG35727
ID   AAG35727 standard; Protein; 326 AA.
XX
AC   AAG35727;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 43687.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   19-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
```

PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.4%; Score 101.5; DB 21; Length 326;
Best Local Similarity 22.1%; Pred. No. 0.044;
Matches 31; Conservative 30; Mismatches 68; Indels 11; Gaps 5;

```
Qy    218 ISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLN 277
          |: |:  :  :     ||    |   |:||  : | :    :|  :    :|:   :
Db     22 intgasslsfvgsafivl---cyclfkelrkfsfklvfylalsdmlcsfflivgdps-kg 77

Qy    278 GICSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTG 337
           ||   ||::||| :|| || : |::   : :   :: |       :: :|| |: ::|
Db     78 ficyaqgytthffcvasflwtttiaftlhrtvvk-hktdvedleamfhlyvwgtslvvt- 135

Qy    338 SIYIVNQIWENDP--SKWNW 355
           ::     |      | |
Db    136 ---virsfgnnhshlgpwcw 152
```

RESULT 13
AAG05052
ID AAG05052 standard; Protein; 331 AA.
XX
AC AAG05052;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1311.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.4%; Score 101.5; DB 21; Length 331;
Best Local Similarity 22.1%; Pred. No. 0.045;
Matches 31; Conservative 30; Mismatches 68; Indels 11; Gaps 5;

```
Qy    218 ISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLN 277
          |: |:  :  :     ||    |   |:||  : | :    :|  :    :|:   :
Db     22 intgasslsfvgsafivl---cyclfkelrkfsfklvfylalsdmlcsfflivgdps-kg 77

Qy    278 GICSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTG 337
```

```
          ||   ||::||| :|| || : |::   : :   :: |       :: :|| |: ::|
Db     78 ficyaqgytthffcvasflwtttiaftlhrtvvk-hktdvedleamfhlyvwgtslvvt- 135

Qy    338 SIYIVNQIWENDP--SKWNW 355
          ::     |      | |
Db    136 ---virsfgnnhshlgpwcw 152


RESULT  14
AAG05051
ID   AAG05051 standard; Protein; 350 AA.
XX
AC   AAG05051;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 1310.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
```

```
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             4.4%;  Score 101.5;  DB 21;  Length 350;
  Best Local Similarity   22.1%;  Pred. No. 0.049;
  Matches   31;  Conservative   30;  Mismatches   68;  Indels   11;  Gaps    5;

Qy    218 ISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLN 277
          |: |:  :  :    ||    |   |:||  : | :   :|  :    :|:   :
Db     41 intgasslsfvgsafivl---cyclfkelrkfsfklvfylalsdmlcsfflivgdps-kg 96
```

```
Qy    278 GICSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTG 337
          ||   ||::||| :|| || : |::  : :   :: |       :: :|| |: ::|
Db     97 ficyaqgytthffcvasflwtttiaftlhrtvvk-hktdvedleamfhlyvwgtslvvt- 154

Qy    338 SIYIVNQIWENDP--SKWNW 355
             ::     |      | |
Db    155 ---virsfgnnhshlgpwcw 171


RESULT  15
AAG35726
ID   AAG35726 standard; Protein; 354 AA.
XX
AC   AAG35726;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 43686.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
```

PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.4%; Score 101.5; DB 21; Length 354;
Best Local Similarity 22.1%; Pred. No. 0.05;
Matches 31; Conservative 30; Mismatches 68; Indels 11; Gaps 5;

Qy 218 ISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLN 277
Db 50 intgasslsfvgsafivl---cyclfkelrkfsfklvfylalsdmlcsfflivgdps-kg 105

```
Qy     278 GICSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTG 337
           ||   ||::||| :|| || : |::   : :   :: |       :: :|| |: ::|
Db     106 ficyaqgytthffcvasflwtttiaftlhrtvvk-hktdvedleamfhlyvwgtslvvt- 163

Qy     338 SIYIVNQIWENDP--SKWNW 355
             ::     |      | |
Db     164 ---virsfgnnhshlgpwcw 180


Search completed: October 30, 2001, 15:19:38
Job time: 2524 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 14:38:59 ; Search time 51.48 Seconds
(without alignments)
170.386 Million cell updates/sec

Title: US-09-619-049-855
Perfect score: 2324
Sequence: 1 MRLPWVIFCTVLLLIFTNNS.....NKFTNEEEGRINCINFDSQT 426

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	118	5.1	884	2	US-08-465-976A-2	Sequence 2, Appli
2	118	5.1	884	2	US-08-982-412-2	Sequence 2, Appli
3	106	4.6	652	3	US-09-110-116-1	Sequence 1, Appli
4	98.5	4.2	256	4	US-08-804-166-2	Sequence 2, Appli
5	98.5	4.2	256	4	US-08-910-991-2	Sequence 2, Appli
6	97.5	4.2	336	4	US-08-804-166-8	Sequence 8, Appli
7	97.5	4.2	336	4	US-08-910-991-8	Sequence 8, Appli
8	97	4.2	344	3	US-09-110-116-4	Sequence 4, Appli
9	97	4.2	611	2	US-08-677-049-2	Sequence 2, Appli
10	96	4.1	886	3	US-09-110-116-3	Sequence 3, Appli
11	92.5	4.0	446	2	US-08-833-610-5	Sequence 5, Appli
12	92.5	4.0	446	3	US-08-834-033A-15	Sequence 15, Appl
13	92.5	4.0	448	1	US-08-366-779-5	Sequence 5, Appli
14	92.5	4.0	448	1	US-08-789-936-5	Sequence 5, Appli
15	91	3.9	307	4	US-08-804-166-4	Sequence 4, Appli
16	91	3.9	307	4	US-08-910-991-4	Sequence 4, Appli
17	90	3.9	280	3	US-08-974-022-46	Sequence 46, Appl
18	90	3.9	455	1	US-08-050-319B-25	Sequence 25, Appl
19	90	3.9	455	1	US-08-321-668-2	Sequence 2, Appli
20	90	3.9	455	1	US-08-837-941-2	Sequence 2, Appli
21	90	3.9	455	2	US-08-126-016-2	Sequence 2, Appli
22	90	3.9	455	2	US-08-465-982-25	Sequence 25, Appl
23	90	3.9	455	4	US-08-815-469-5	Sequence 5, Appli
24	88	3.8	553	3	US-08-845-546-12	Sequence 12, Appl
25	86	3.7	285	4	US-08-804-166-6	Sequence 6, Appli
26	86	3.7	285	4	US-08-910-991-6	Sequence 6, Appli
27	83.5	3.6	359	2	US-08-748-485-5	Sequence 5, Appli
28	83	3.6	199	1	US-08-050-319B-48	Sequence 48, Appl
29	83	3.6	199	2	US-08-465-982-48	Sequence 48, Appl
30	82.5	3.5	418	1	US-07-946-232-8	Sequence 8, Appli
31	82.5	3.5	423	1	US-08-073-799C-8	Sequence 8, Appli
32	82.5	3.5	423	1	US-07-947-672-8	Sequence 8, Appli
33	82.5	3.5	423	1	US-08-432-043-8	Sequence 8, Appli
34	82.5	3.5	423	2	US-08-660-963-8	Sequence 8, Appli
35	82.5	3.5	471	1	US-07-817-920-8	Sequence 8, Appli
36	82.5	3.5	471	1	US-08-370-542-7	Sequence 7, Appli
37	82.5	3.5	471	1	US-08-117-006-8	Sequence 8, Appli
38	82.5	3.5	471	1	US-08-216-594-8	Sequence 8, Appli
39	82.5	3.5	471	1	US-08-542-358-7	Sequence 7, Appli
40	82.5	3.5	471	2	US-08-244-434-2	Sequence 2, Appli
41	82.5	3.5	471	3	US-09-018-351-7	Sequence 7, Appli
42	82.5	3.5	471	4	US-09-032-742-8	Sequence 8, Appli
43	82.5	3.5	471	5	PCT-US93-00149-8	Sequence 8, Appli
44	82.5	3.5	478	3	US-09-292-071-31	Sequence 31, Appl
45	82.5	3.5	478	3	US-09-292-071-33	Sequence 33, Appl

ALIGNMENTS

```
RESULT 1
US-08-465-976A-2
; Sequence 2, Application US/08465976A
; Patent No. 5869632
;  GENERAL INFORMATION:
;    APPLICANT:  SOPPET, DANIEL R
;    APPLICANT:  LI, YI
;    APPLICANT:  ROSEN, CRAIG A
;    APPLICANT:  RUBEN, STEVEN M
;    TITLE OF INVENTION:  HUMAN G-PROTEIN RECEPTOR
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN GILFILLAN, CECCHI
;      ADDRESSEE:  STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NJ
;      COUNTRY:  US
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,976A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FERRARO, GREGORY F
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-444
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201) 994-1700
;      TELEFAX:  (201) 994-1744
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  884 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-465-976A-2

 Query Match              5.1%;  Score 118;  DB 2;  Length 884;
 Best Local Similarity   22.1%;  Pred. No. 0.0021;
 Matches   50;  Conservative   33;  Mismatches   93;  Indels   50;  Gaps    9;

Qy    207 MWIIR-----HRCTSH----------------ISP-----GSLEIL--------IITMIC 232
          :|  |     ||  ||                ||     | ||:|         :::
```

```
Db     56 VWTARDCELVHRNGSHARCRCSRTGTFGVLMDASPRERLEGDLELLAVFTHVVVAVSVAA 115

Qy    233 FVLTIAVYLYIKKLR-NVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFR 291
          ||| |: | :: |: || |           : | ::  |      :|:      |:|
Db    116 LVLTAAILLSLRSLKSNVRGIHANVAAALGVAELLFLLGIHRTHNQLVCTAVAILLHYFF 175

Qy    292 MASNLWLSVISYHTWKVLTSLNRVDPN-YRFLRYNAFVWSTAAIMTGSIYIVNQIWENDP 350
          :::  || |   | :::      ||    ||  |:|  |   |:: |    ::     :|
Db    176 LSTFAWLFVQGLHLYRMQVEPRNVDRGAMRF--YHALGWGVPAVLLGLAVGLDPEGYGNP 233

Qy    351 SKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTA 396
                    | :   | :| : :|| : :   |  || | |
Db    234 D----------FCWISVHEPLIWSF-AGPVVLVIVMNGTMFLLAA 267


RESULT   2
US-08-982-412-2
; Sequence 2, Application US/08982412
; Patent No. 5958729
;  GENERAL INFORMATION:
;    APPLICANT:  SOPPET, DANIEL R
;    APPLICANT:  LI, YI
;    APPLICANT:  ROSEN, CRAIG A
;    APPLICANT:  RUBEN, STEVEN M
;    TITLE OF INVENTION:  HUMAN G-PROTEIN RECEPTOR
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HUMAN GENOME SCIENCES, INC.
;      STREET:  9410 KEY WEST AVENUE
;      CITY:  ROCKVILLE,
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/982,412
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROOKES, ANDERS A
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF181PCT2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8439
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  884 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-982-412-2


  Query Match             5.1%;  Score 118;  DB 2;  Length 884;
  Best Local Similarity   22.1%;  Pred. No. 0.0021;
  Matches   50;  Conservative   33;  Mismatches   93;  Indels   50;  Gaps    9;

Qy    207 MWIIR-----HRCTSH----------------ISP-----GSLEIL--------IITMIC 232
          :|  |     ||  ||                ||     | ||:|          :::
Db     56 VWTARDCELVHRNGSHARCRCSRTGTFGVLMDASPRERLEGDLELLAVFTHVVVAVSVAA 115

Qy    233 FVLTIAVYLYIKKLR-NVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFR 291
          ||| |: | :: |: || |           : | ::  |      :|:      |:|
Db    116 LVLTAAILLSLRSLKSNVRGIHANVAAALGVAELLFLLGIHRTHNQLVCTAVAILLHYFF 175

Qy    292 MASNLWLSVISYHTWKVLTSLNRVDPN-YRFLRYNAFVWSTAAIMTGSIYIVNQIWENDP 350
          :::  || |   | :::      ||    ||  |:|  |   |:: |    ::     :|
Db    176 LSTFAWLFVQGLHLYRMQVEPRNVDRGAMRF--YHALGWGVPAVLLGLAVGLDPEGYGNP 233

Qy    351 SKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTA 396
                    | :   | :| : :|| : :   |  || | |
Db    234 D----------FCWISVHEPLIWSF-AGPVVLVIVMNGTMFLLAA 267


RESULT   3
US-09-110-116-1
; Sequence 1, Application US/09110116
; Patent No. 6013479
; GENERAL INFORMATION:
;  APPLICANT: Xu, Hong
;  APPLICANT: Cohan, Victoria L.
;  APPLICANT: Stuart, Susan G.
;  TITLE OF INVENTION: HUMAN EMR1-LIKE G PROTEIN COUPLED
;  TITLE OF INVENTION: RECEPTOR
;  FILE REFERENCE: PF-0550 US
;  CURRENT APPLICATION NUMBER: US/09/110,116
;  CURRENT FILING DATE: 1998-07-02
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 652
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
;   FEATURE:
;   OTHER INFORMATION: 429905, EOSINOT03
US-09-110-116-1


  Query Match             4.6%;  Score 106;  DB 3;  Length 652;
  Best Local Similarity   20.0%;  Pred. No. 0.023;
  Matches   54;  Conservative   41;  Mismatches   93;  Indels   82;  Gaps   13;

Qy    173 WFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEI-----LI 227
          | :|| | : |:   :    |   |        | : :    ||      | :     |
Db    316 WSRDGCFLIHVN---KSHTMCNCSH-------LSSFAVLMALTSQEEDPVLTVITYVGLS 365

Qy    228 ITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQ-CLIMILDHLNLLNGI------- 279
          ::::| :|    :|  | ::| :          |  :|  | : | ||  | ||
Db    366 VSLLCLLLAALTFLLCKAIQNTS--------TSLHLQLSLCLFLAHLLFLVGIDRTEPKV 417

Qy    280 -CSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGS 338
           ||    : |:  :|:  |: :   | :  ||: |    ||  :
Db    418 LCSIIAGALHYLYLAAFTWMLLEGVHLF--LTARNLTVVNYSSI---------------- 459

Qy    339 IYIVNQIWENDPSKWNWLPLVGF------IRCSVKDWHPSV-------WIYIS------- 378
                  |   ||   | ||:      :  |   | | :       |:::
Db    460 ---------NRLMKWIMFP-VGYGVPAVTVAISAASW-PHLYGTADRCWLHLDQGFMWSF 508

Qy    379 -GPSLALSTFNVAMFALTAIYIRKVKGGIN 407
           ||  |: : |: :| |    :::    :|
Db    509 LGPVCAIFSANLVLFILVFWILKRKLSSLN 538


RESULT   4
US-08-804-166-2
; Sequence 2, Application US/08804166
; Patent No. 6193972
;  GENERAL INFORMATION:
;    APPLICANT:  Campbell, Robert K.
;    APPLICANT:  Jameson, Bradford A.
;    APPLICANT:  Chappel, Scott C.
;    TITLE OF INVENTION:  HYBRID PROTEINS
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street N.W., Ste. 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  22207
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,166
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/011,936
;      FILING DATE:  20 February 1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Browdy, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  CAMPBELL=2A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  256 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-804-166-2


  Query Match             4.2%;  Score 98.5;  DB 4;  Length 256;
  Best Local Similarity   23.0%;  Pred. No. 0.036;
  Matches   61;  Conservative   30;  Mismatches   81;  Indels   93;  Gaps   16;

Qy      37 YIERQNDS----------YLYDDIEIPASLTGYYEFRQFGDGSITPIEKHLRACV-CSVR 85
           ||  ||:|          |||:|   |   |   : |:   || |  | ||| |: ||
Db      31 YIHPQNNSICCTKCHKGTYLYNDCPGPGQDT---DCRECESGSFTASENHLRHCLSCS-- 85

Qy      86 PCIRICCPAKNFLANGKCDDGLKEELARFKPYIYFTYMDLQARVPLTDMAIIRDEFFDCD 145
                      ||    ::|:              :| ::   : ||    |
Db      86 ---------------KC----RKEM--------------GQVEISSCTVDRDTVCGCR 110

Qy     146 EMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFWVTVDL-FMEKQDYCLYRHNFDSDFP 204
           :    : : ::  |   | || | | :   :|    || |   |||:     |
Db     111 K----NQYRHYWSENLFQCFN-CSLCL---NG----TVHLSCQEKQNTVCTCH------- 151

Qy     205 KSMWIIRHRCTS--HISPGSLE--------------ILIITMICFVLTIAVYLYIKKL- 246
             ::  : | |    :||  |                ||     ||        |  ||
Db     152 AGFFLRENECVSCAGAAPGCPECTLQENPFFSQPGAPILQCMGCCFSRAYPTPLRSKKTM 211

Qy     247 ---RNVTGKCIVCCIVSRFIQCLIM 268
              :||| :   ||:   : :  :|
Db     212 LVQKNVTSES-TCCVAKSYNRVTVM 235


RESULT   5
US-08-910-991-2
; Sequence 2, Application US/08910991
; Patent No. 6194177
;  GENERAL INFORMATION:
;    APPLICANT:  Campbell, Robert K.
;    APPLICANT:  Jameson, Bradford A.
;    APPLICANT:  Chappel, Scott C.
;    TITLE OF INVENTION:  HYBRID PROTEINS
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street N.W., Ste. 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  22207
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/910,991
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/804,166
;      FILING DATE:  20 February 1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/011,936
;      FILING DATE:  20 February 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  YUN, Allen C.
;      REGISTRATION NUMBER:  37,971
;      REFERENCE/DOCKET NUMBER:  CAMPBELL=2B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  256 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-910-991-2


  Query Match             4.2%;  Score 98.5;  DB 4;  Length 256;
  Best Local Similarity   23.0%;  Pred. No. 0.036;
  Matches   61;  Conservative   30;  Mismatches   81;  Indels   93;  Gaps   16;

Qy      37 YIERQNDS----------YLYDDIEIPASLTGYYEFRQFGDGSITPIEKHLRACV-CSVR 85
           ||  ||:|          |||:|   |   |   : |:   || |  | ||| |: ||
Db      31 YIHPQNNSICCTKCHKGTYLYNDCPGPGQDT---DCRECESGSFTASENHLRHCLSCS-- 85

Qy      86 PCIRICCPAKNFLANGKCDDGLKEELARFKPYIYFTYMDLQARVPLTDMAIIRDEFFDCD 145
                      ||    ::|:              :| ::   : ||    |
Db      86 ---------------KC----RKEM--------------GQVEISSCTVDRDTVCGCR 110

Qy     146 EMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFWVTVDL-FMEKQDYCLYRHNFDSDFP 204
           :    : : ::  |   | || | | :   :|    || |   |||:     |
Db     111 K----NQYRHYWSENLFQCFN-CSLCL---NG----TVHLSCQEKQNTVCTCH------- 151

Qy     205 KSMWIIRHRCTS--HISPGSLE--------------ILIITMICFVLTIAVYLYIKKL- 246
             ::  : | |    :||  |                ||     ||        |  ||
Db     152 AGFFLRENECVSCAGAAPGCPECTLQENPFFSQPGAPILQCMGCCFSRAYPTPLRSKKTM 211

Qy     247 ---RNVTGKCIVCCIVSRFIQCLIM 268
              :||| :   ||:   : :  :|
Db     212 LVQKNVTSES-TCCVAKSYNRVTVM 235


RESULT   6
US-08-804-166-8
; Sequence 8, Application US/08804166
; Patent No. 6193972
;  GENERAL INFORMATION:
;    APPLICANT:  Campbell, Robert K.
;    APPLICANT:  Jameson, Bradford A.
;    APPLICANT:  Chappel, Scott C.
;    TITLE OF INVENTION:  HYBRID PROTEINS
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street N.W., Ste. 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  22207
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/804,166
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/011,936
;       FILING DATE:  20 February 1996
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Browdy, Roger L.
;       REGISTRATION NUMBER:  25,618
;       REFERENCE/DOCKET NUMBER:  CAMPBELL=2A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 628-5197
;       TELEFAX:  (202) 737-3528
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  336 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-804-166-8


  Query Match             4.2%;  Score 97.5;  DB 4;  Length 336;
  Best Local Similarity   19.0%;  Pred. No. 0.066;
  Matches   64;  Conservative   38;  Mismatches   78;  Indels  157;  Gaps   17;

Qy      37 YIERQNDS----------YLYDDIEIPASLTGYYEFRQFGDGSITPIEKHLRACV-CSVR 85
           ||  ||:|          |||:|   |   |   : |:   || |  | ||| |: ||
Db      31 YIHPQNNSICCTKCHKGTYLYNDCPGPGQDT---DCRECESGSFTASENHLRHCLSCS-- 85

Qy      86 PCIRICCPAKNFLANGKCDDGLKEELARFKPYIYFTYMDLQARVPLTDMAIIRDEFFDCD 145
                          ||    ::|:                :| ::   : ||    |
Db      86 ---------------KC----RKEM--------------GQVEISSCTVDRDTVCGCR 110

Qy     146 EMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFWVTVDL-FMEKQDYCLYRHNFDSDFP 204
           :   : : ::  |   | || | | :   :|    || |   |||:
Db     111 K----NQYRHYWSENLFQCFN-CSLCL---NG----TVHLSCQEKQNTV----------- 147

Qy     205 KSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQ 264
                 || |                          ::::      :|: |    : ::
Db     148 ---------CTCHAG--------------------FFLRE-----NECVSCSNCKKSLE 172

Qy     265 CLIMILDHLNLLNGI------------CSP----------------------AGYSSHF 289
           |  : |  :  : |              | |                        |||
Db     173 CTKLCLPQIENVKGTEDSGTTAGAGPRCRPINATLAVEKEGCPVCITVNTTICAGYCPTM 232

Qy     290 FRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNA 326
           |:                || :| :|  ||| :|: :
Db     233 TRVLQG------------VLPALPQVVCNYRDVRFES 257


RESULT   7
US-08-910-991-8
; Sequence 8, Application US/08910991
; Patent No. 6194177
;  GENERAL INFORMATION:
;     APPLICANT:  Campbell, Robert K.
;     APPLICANT:  Jameson, Bradford A.
;     APPLICANT:  Chappel, Scott C.
;     TITLE OF INVENTION:  HYBRID PROTEINS
;     NUMBER OF SEQUENCES:  22
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  BROWDY AND NEIMARK
;       STREET:  419 Seventh Street N.W., Ste. 300
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  USA
;       ZIP:  22207
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/910,991
;       FILING DATE:
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/804,166
;       FILING DATE:  20 February 1997
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/011,936
;       FILING DATE:  20 February 1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  YUN, Allen C.
;       REGISTRATION NUMBER:  37,971
;       REFERENCE/DOCKET NUMBER:  CAMPBELL=2B
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 628-5197
;       TELEFAX:  (202) 737-3528
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  336 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-910-991-8


  Query Match             4.2%;  Score 97.5;  DB 4;  Length 336;
  Best Local Similarity   19.0%;  Pred. No. 0.066;
  Matches   64;  Conservative   38;  Mismatches   78;  Indels  157;  Gaps   17;

Qy      37 YIERQNDS----------YLYDDIEIPASLTGYYEFRQFGDGSITPIEKHLRACV-CSVR 85
           ||  ||:|          |||:|   |   |   : |:   || |  | ||| |: ||
Db      31 YIHPQNNSICCTKCHKGTYLYNDCPGPGQDT---DCRECESGSFTASENHLRHCLSCS-- 85

Qy      86 PCIRICCPAKNFLANGKCDDGLKEELARFKPYIYFTYMDLQARVPLTDMAIIRDEFFDCD 145
                          ||    ::|:                :| ::   : ||    |
Db      86 ---------------KC----RKEM--------------GQVEISSCTVDRDTVCGCR 110

Qy     146 EMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFWVTVDL-FMEKQDYCLYRHNFDSDFP 204
           :   : : ::  |   | || | | :   :|    || |   |||:
Db     111 K----NQYRHYWSENLFQCFN-CSLCL---NG----TVHLSCQEKQNTV----------- 147

Qy     205 KSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQ 264
                 || |                          ::::      :|: |    : ::
Db     148 ---------CTCHAG--------------------FFLRE-----NECVSCSNCKKSLE 172

Qy     265 CLIMILDHLNLLNGI------------CSP----------------------AGYSSHF 289
           |  : |  :  : |              | |                        |||
Db     173 CTKLCLPQIENVKGTEDSGTTAGAGPRCRPINATLAVEKEGCPVCITVNTTICAGYCPTM 232

Qy     290 FRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNA 326
           |:                || :| :|  ||| :|: :
Db     233 TRVLQG------------VLPALPQVVCNYRDVRFES 257


RESULT   8
US-09-110-116-4
; Sequence 4, Application US/09110116
; Patent No. 6013479
; GENERAL INFORMATION:
;  APPLICANT: Xu, Hong
;  APPLICANT: Cohan, Victoria L.
;  APPLICANT: Stuart, Susan G.
;  TITLE OF INVENTION: HUMAN EMR1-LIKE G PROTEIN COUPLED
;  TITLE OF INVENTION: RECEPTOR
;  FILE REFERENCE: PF-0550 US
;  CURRENT APPLICATION NUMBER: US/09/110,116
;  CURRENT FILING DATE: 1998-07-02
;  NUMBER OF SEQ ID NOS: 4
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
;   FEATURE:
;   OTHER INFORMATION: 2935597, GenBank
US-09-110-116-4


  Query Match             4.2%;  Score 97;  DB 3;  Length 344;
  Best Local Similarity   22.7%;  Pred. No. 0.076;
  Matches   54;  Conservative   38;  Mismatches   70;  Indels   76;  Gaps   15;

Qy     213 RCTSHISPGSLEILI-------------------ITMICFVLTIAVYLYIKKLRNVTGK 252
           ||| |:|  |  :|:                    ::::| :|    :|  | ::| :
Db      35 RCT-HLS--SFAVLMAHYDVQEDPVLTVITYMGLSVSLLCLLLAALTFLLCKAIQNTS-- 89

Qy     253 CIVCCIVSRFIQ-CLIMILDHLNLLNGI--------CSPAGYSSHFFRMASNLWLSVISY 303
               |  :|  | : | ||  |  |          ||    : |:  :|:  |: : :
Db      90 ------TSLHLQLSLCLFLAHLLFLVAIDQTGHKVLCSIIAGTLHYLYLATFTWMLLEAL 143

Qy     304 HTWKVLTSLNRVDPNY----RFLRYNAFVWS------TAAIMTGS---IYIVNQIWENDP 350
           : :  ||: |    ||    ||::   |          | ||   |   :|         |
Db     144 YLF--LTARNLTVVNYSSINRFMKKLMFPVGYGVPAVTVAISAASRPHLY-------GTP 194

Qy     351 SKWNWL-PLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGIN 407
           |:  || |  |||           | :: ||  |: : |: :| :|   ::      :|
Db     195 SRC-WLQPEKGFI----------WGFL-GPVCAIFSVNLVLFLVTLWILKNRLSSLN 239


RESULT   9
US-08-677-049-2
; Sequence 2, Application US/08677049
; Patent No. 5858707
;  GENERAL INFORMATION:
;    APPLICANT:  Guimaraes, M. Jorge
;    APPLICANT:  Bazan, J. Fernando
;    APPLICANT:  McClanahan, Terrill K.
;    APPLICANT:  Zlotnik, Albert
;    TITLE OF INVENTION:  PURIFIED MAMMALIAN NUCLEOBASE PERMEASES;
;    TITLE OF INVENTION:  NUCLEIC ACIDS; ANTIBODIES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/677,049
;      FILING DATE:  03-JUL-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/000,788
;      FILING DATE:  03-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
;      REFERENCE/DOCKET NUMBER:  DX0511
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-852-9196
;      TELEFAX:  415-496-1200
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 611 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-677-049-2


  Query Match             4.2%;  Score 97;  DB 2;  Length 611;
  Best Local Similarity   25.3%;  Pred. No. 0.17;
  Matches   58;  Conservative   28;  Mismatches   75;  Indels   68;  Gaps   13;

Qy     219 SPG----SLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLN 274
           :||     || :  :|  |  |      |:   || |:|      :||  :|  | :|
Db     139 TPGNASLSLPLCSLTRSCHGLE----LWNTSLREVSG----AVVVSGLLQGTIGLLGVPG 190

Qy     275 LLNGICSP---------AGYSSH--FFRMASNLW---------LSVISYH---------T 305
             :   | |         || |:|    :  |  |         : | | |         :
Db     191 RVFPYCGPLVLAPSLVVAGLSAHKEVAQFCSAHWGLALLLILLMVVCSQHLGSCQIPLCS 250

Qy     306 WK-VLTSLNRVDPNYRFLRY---NAFVWSTAAIMTGSIYIVNQIWENDPSKWNWLPLVGF 361
           |:   || :   | :| |       | ||  :| : |:   |  |: |   : | |||
Db     251 WRPSSTSTHICIPVFRLLSVLAPVACVWFISAFV-GTSVIPLQLSEPSDAPWFWLP---- 305

Qy     362 IRCSVKDWHPSVWIY--------ISGPSLAL--STFNVAMFALTAIYIR 400
                 ||  | :          :| |:||  || ::  :||    :|
Db     306 --------HPGEWEWPLLTPRALAAGISMALAASTSSLGCYALCGQLLR 346


RESULT  10
US-09-110-116-3
; Sequence 3, Application US/09110116
; Patent No. 6013479
; GENERAL INFORMATION:
;  APPLICANT: Xu, Hong
;  APPLICANT: Cohan, Victoria L.
;  APPLICANT: Stuart, Susan G.
;  TITLE OF INVENTION: HUMAN EMR1-LIKE G PROTEIN COUPLED
;  TITLE OF INVENTION: RECEPTOR
;  FILE REFERENCE: PF-0550 US
;  CURRENT APPLICATION NUMBER: US/09/110,116
;  CURRENT FILING DATE: 1998-07-02
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 886
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
;   FEATURE:
;   OTHER INFORMATION: 784994, GenBank
US-09-110-116-3


  Query Match             4.1%;  Score 96;  DB 3;  Length 886;
  Best Local Similarity   21.0%;  Pred. No. 0.36;
  Matches   48;  Conservative   40;  Mismatches   75;  Indels   66;  Gaps   11;

Qy     202 DFPKSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSR 261
           ||  |::||     ||:       :||:::| || || :|  : :||           :
Db     599 DF--SLYII-----SHVG------IIISLVCLVLAIATFLLCRSIRN----------HNT 635

Qy     262 FIQ---CLIMILDHLNLLNGI-------CSPAGYSSHFFRMASNLWLSVISYHTWKVLT 310
           ::    |: ::|     | ||        |:      |:  :|   |: | :   : ::
Db     636 YLHLHLCVCLLLAKTLFLAGIHKTDNKTGCAIIAGFLHYLFLACFFWMLVEAVILFLMVR 695

Qy     311 SLNRVD----PNYRFLRYNAF--------VWSTAAIMTGSIYIVNQIWENDPSKWNWLPL 358
           :|  |:     | : |   ||         |  :|::      : |: | |  :
Db     696 NLKVVNYFSSRNIKMLHICAFGYGLPMLVVVISASVQPQGYGMHNRCWLNTET------- 748

Qy     359 VGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGIN 407
             |||           | :: ||   :   |  :   |   :|:    :|
Db     749 -GFI----------WSFL-GPVCTVIVINSLLLTWTLWILRQRLSSVN 784


RESULT  11
US-08-833-610-5
```

```
; Sequence 5, Application US/08833610
; Patent No. 5972664
;  GENERAL INFORMATION:
;    APPLICANT:  KNUTZON, DEBORAH
;    APPLICANT:  MURKERJI, PRADIP
;    APPLICANT:  HUANG, YUNG-SHENG
;    APPLICANT:  THURMOND, JENNIFER
;    APPLICANT:  CHAUDHARY, SUNITA
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR SYNTHESIS
;    TITLE OF INVENTION:  OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  RAE-VENTER LAW GROUP, P.C.
;      STREET:  260 SHERIDAN AVE, P.O. BOX 60039
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Relèase #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/833,610
;      FILING DATE:  11-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  RAE-VENTER, BARBARA
;      REGISTRATION NUMBER:  32,750
;      REFERENCE/DOCKET NUMBER:  CGNE.123.00US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)328-4400
;      TELEFAX:  (650)328-4477
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  446 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-833-610-5


  Query Match             4.0%;  Score 92.5;  DB 2;  Length 446;
  Best Local Similarity   19.5%;  Pred. No. 0.32;
  Matches   59;  Conservative   43;  Mismatches   82;  Indels  119;  Gaps   15;

Qy    156 FLEEVSIQIFNKCGL-IVWFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRC 214
          | | | : :|: | :  :| | |  |:        |    |    ||   | | |    |
Db    133 FCEGVLVHLFSGCLMGFLWIQSG--WI-------GHDAGHYMVVSDSRLNKFMGIFAANC 183

Qy    215 TS--------------HISPGSLEI--------LIITMICFVLTIAVYLYIKKL------ 246
           |              ||:  |||           ::    |  ::  : | |:|
Db    184 LSGISIGWWKWNHNAHHIACNSLEYDPDLQYIPFLVVSSKFFGSLTSHFYEKRLTFDSLS 243

Qy    247 ------RNVTGKCIVCCI-VSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLS 299
                :: |   |:|   :: ::| |||:|   |
Db    244 RFFVSYQHWTFYPIMCAARLNMYVQSLIMLLTKRN------------------------- 278

Qy    300 VISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWENDPSKWNWLPLV 359
           :||   ::|  |                |:|                       | ||:
Db    279 -VSYRAQELLGCL--------------VFSI----------------------WYPLL 299

Qy    360 GFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFAL----TAIYIRKVKGGINKFTNEEEG 415
            : | : :|   :   |:  ||:::      |:|    :::|: | ||  | |  : :|
Db    300 --VSC-LPNWGERIMFVIA--SLSVTGMQQVQFSLNHFSSSVYVGKPKGN-NWFEKQTDG 353

Qy    416 RIN 418
          ::
Db    354 TLD 356
```

```
RESULT  12
US-08-834-033A-15
; Sequence 15, Application US/08834033A
; Patent No. 6075183
;  GENERAL INFORMATION:
;    APPLICANT:  KNUTZON, DEBORAH
;    APPLICANT:  MUKERJI, PRADIP
;    APPLICANT:  HUANG, YUNG-SHENG
;    APPLICANT:  THURMOND, JENNIFER
;    APPLICANT:  CHAUDHARY, SUNITA
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR SYNTHESIS
;    TITLE OF INVENTION:  OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LIMBACH AND LIMBACH, L.L.P.
;      STREET:  2001 FERRY BUILDING
;      CITY:  SAN FRANCISCO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,033A
;      FILING DATE:  11-APR-1997
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WARD, MICHAEL R.
;      REGISTRATION NUMBER:  38,651
;      REFERENCE/DOCKET NUMBER:  CGAB-300.USA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 433-4150
;      TELEFAX:  (415) 433-8716
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  446 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-834-033A-15


  Query Match             4.0%;  Score 92.5;  DB 3;  Length 446;
  Best Local Similarity   19.5%;  Pred. No. 0.32;
  Matches   59;  Conservative   43;  Mismatches   82;  Indels  119;  Gaps   15;

Qy    156 FLEEVSIQIFNKCGL-IVWFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRC 214
          | | | : :|: | :  :| | |  |:        |    |    ||   | | |    |
Db    133 FCEGVLVHLFSGCLMGFLWIQSG--WI-------GHDAGHYMVVSDSRLNKFMGIFAANC 183

Qy    215 TS--------------HISPGSLEI--------LIITMICFVLTIAVYLYIKKL------ 246
           |              ||:  |||           ::    |  ::  : | |:|
Db    184 LSGISIGWWKWNHNAHHIACNSLEYDPDLQYIPFLVVSSKFFGSLTSHFYEKRLTFDSLS 243

Qy    247 ------RNVTGKCIVCCI-VSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLS 299
                :: |   |:|   :: ::| |||:|   |
Db    244 RFFVSYQHWTFYPIMCAARLNMYVQSLIMLLTKRN------------------------- 278

Qy    300 VISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWENDPSKWNWLPLV 359
           :||   ::|  |                |:|                       | ||:
Db    279 -VSYRAQELLGCL--------------VFSI----------------------WYPLL 299

Qy    360 GFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFAL----TAIYIRKVKGGINKFTNEEEG 415
            : | : :|   :   |:  ||:::      |:|    :::|: | ||  | |  : :|
Db    300 --VSC-LPNWGERIMFVIA--SLSVTGMQQVQFSLNHFSSSVYVGKPKGN-NWFEKQTDG 353

Qy    416 RIN 418
```

```
Db    354 TLD 356


RESULT  13
US-08-366-779-5
; Sequence 5, Application US/08366779
; Patent No. 5614393
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Terry L.
;    APPLICANT:  Reddy, Avutu S.
;    APPLICANT:  Nuccio, Michael
;    APPLICANT:  Freyssinet, Georges L.
;    APPLICANT:  Nunberg, Andrew N.
;    TITLE OF INVENTION:  PRODUCTION OF GAMMA LINOLENIC ACID BY A
;    TITLE OF INVENTION:  DELTA 6-DESATURASE
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  United States
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/366,779
;      FILING DATE:  30-DEC-1994
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Presser, Leopold
;      REGISTRATION NUMBER:  19,827
;      REFERENCE/DOCKET NUMBER:  8383ZYXW
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  448 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-366-779-5


  Query Match             4.0%;  Score 92.5;  DB 1;  Length 448;
  Best Local Similarity   19.5%;  Pred. No. 0.32;
  Matches   59;  Conservative   43;  Mismatches   82;  Indels  119;  Gaps   15;

Qy    156 FLEEVSIQIFNKCGL-IVWFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRC 214
          | | | : :|: | :  :| | |  |:         |    |    ||   | | |    |
Db    133 FCEGVLVHLFSGCLMGFLWIQSG--WI-------GHDAGHYMVVSDSRLNKFMGIFAANC 183

Qy    215 TS--------------HISPGSLEI--------LIITMICFVLTIAVYLYIKKL------ 246
            |              ||:  |||          ::    |  ::  : | |:|
Db    184 LSGISIGWWKWNHNAHHIACNSLEYDPDLQYIPFLVVSSKFFGSLTSHFYEKRLTFDSLS 243

Qy    247 ------RNVTGKCIVCCI-VSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLS 299
                :: |   |:|   :: ::| |||:|   |
Db    244 RFFVSYQHWTFYPIMCAARLNMYVQSLIMLLTKRN------------------------ 278

Qy    300 VISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWENDPSKWNWLPLV 359
           :||   ::|  |              |:|                          | ||:
Db    279 -VSYRAQELLGCL--------------VFSI----------------------WYPLL 299

Qy    360 GFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFAL----TAIYIRKVKGGINKFTNEEEG 415
           : | : :|   :   |:  ||:::      |:|    :::|: | ||  | |  : :|
Db    300 --VSC-LPNWGERIMFVIA--SLSVTGMQQVQFSLNHFSSSVYVGKPKGN-NWFEKQTDG 353

Qy    416 RIN 418
          ::
Db    354 TLD 356


RESULT  14
US-08-789-936-5
; Sequence 5, Application US/08789936
; Patent No. 5789220
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Terry L.
;    APPLICANT:  Reddy, Avutu S.
;    APPLICANT:  Nuccio, Michael
;    APPLICANT:  Freyssinet, Georges L.
;    APPLICANT:  Nunberg, Andrew N.
;    TITLE OF INVENTION:  PRODUCTION OF GAMMA LINOLENIC ACID BY A
;    TITLE OF INVENTION:  DELTA 6-DESATURASE
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  United States
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/789,936
;      FILING DATE:  28-JAN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/366,779
;      FILING DATE:  30-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Presser, Leopold
;      REGISTRATION NUMBER:  19,827
;      REFERENCE/DOCKET NUMBER:  8383ZYXW
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  448 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
US-08-789-936-5


  Query Match             4.0%;  Score 92.5;  DB 1;  Length 448;
  Best Local Similarity   19.5%;  Pred. No. 0.32;
  Matches   59;  Conservative   43;  Mismatches   82;  Indels  119;  Gaps   15;

Qy    156 FLEEVSIQIFNKCGL-IVWFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRC 214
          | | | : :|: | :  :| | |  |:         |    |    ||   | | |    |
Db    133 FCEGVLVHLFSGCLMGFLWIQSG--WI-------GHDAGHYMVVSDSRLNKFMGIFAANC 183

Qy    215 TS--------------HISPGSLEI--------LIITMICFVLTIAVYLYIKKL------ 246
            |              ||:  |||          ::    |  ::  : | |:|
Db    184 LSGISIGWWKWNHNAHHIACNSLEYDPDLQYIPFLVVSSKFFGSLTSHFYEKRLTFDSLS 243

Qy    247 ------RNVTGKCIVCCI-VSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLS 299
                :: |   |:|   :: ::| |||:|   |
Db    244 RFFVSYQHWTFYPIMCAARLNMYVQSLIMLLTKRN------------------------ 278

Qy    300 VISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWENDPSKWNWLPLV 359
           :||   ::|  |              |:|                          | ||:
```

```
Db    279 -VSYRAQELLGCL---------------VFSI----------------------WYPLL 299

Qy    360 GFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFAL----TAIYIRKVKGGINKFTNEEEG 415
          : | : :|    :    |:  ||:::      |:|    :::|: | ||  | |  : :|
Db    300 --VSC-LPNWGERIMFVIA--SLSVTGMQQVQFSLNHFSSSVYVGKPKGN-NWFEKQTDG 353

Qy    416 RIN 418
          ::
Db    354 TLD 356


RESULT  15
US-08-804-166-4
; Sequence 4, Application US/08804166
; Patent No. 6193972
;  GENERAL INFORMATION:
;    APPLICANT:  Campbell, Robert K.
;    APPLICANT:  Jameson, Bradford A.
;    APPLICANT:  Chappel, Scott C.
;    TITLE OF INVENTION:  HYBRID PROTEINS
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street N.W., Ste. 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  22207
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,166
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/011,936
;      FILING DATE:  20 February 1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Browdy, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  CAMPBELL=2A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  307 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-804-166-4


  Query Match             3.9%;  Score 91;  DB 4;  Length 307;
  Best Local Similarity   22.2%;  Pred. No. 0.27;
  Matches   67;  Conservative   32;  Mismatches   87;  Indels  116;  Gaps   17;

Qy     37 YIERQNDS----------YLYDDIEIPASLTGYYEFRQFGDGSITPIEKHLRACV-CSVR 85
          ||  ||:|          |||:|   |   |   : |:   || |  | ||| |: ||
Db     31 YIHPQNNSICCTKCHKGTYLYNDCPGPGQDT---DCRECESGSFTASENHLRHCLSCS-- 85

Qy     86 PCIRICCPAKNFLANGKCDDGLKEELARFKPYIYFTYMDLQARVPLTDMAIIRDEFFDCD 145
                       ||    ::|:               :| ::   : ||    |
Db     86 ---------------KC----RKEM--------------GQVEISSCTVDRDTVCGCR 110

Qy    146 EMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFWVTVDL-FMEKQDYCLYRHNFDSDFP 204
          :    : : ::  |   | || | | :   :|    || |   |||:     |
Db    111 K----NQYRHYWSENLFQCFN-CSLCL---NG----TVHLSCQEKQNTVCTCH------- 151

Qy    205 KSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQ 264
             ::  : | |    |                |   | ::|       | ||  |:
Db    152 AGFFLRENECVSCAGAGPR----------CRPINATLAVEK-----EGCPVCITVN---- 192

Qy    265 CLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRY 324
                      ||  |||     |:               || :| :|  ||| :|:
Db    193 -----------TTIC--AGYCPTMTRVLQG-----------VLPALPQVVCNYRDVRF 226

Qy    325 NA 326
          :
Db    227 ES 228


Search completed: October 30, 2001, 15:20:41
Job time: 2502 sec
```

THIS PAGE BLANK (USPTO)

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 15:09:54 ; Search time 55.47 Seconds
(without alignments)
585.007 Million cell updates/sec

Title: US-09-619-049-855
Perfect score: 2324
Sequence: 1 MRLPWVIFCTVLLLIFTNNS.....NKFTNEEEGRINCINFDSQT 426

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	118	5.1	1364	2	T00250	MEGF2 protein - hu
2	116	5.0	1582	2	T15308	hypothetical prote
3	109.5	4.7	564	2	C42523	A55R protein - vac
4	105.5	4.5	323	2	T29673	hypothetical prote
5	105	4.5	1209	2	T16663	hypothetical prote
6	104.5	4.5	564	2	JQ1792	SalF17R protein -
7	102	4.4	3034	2	T14119	seven-pass transme
8	101.5	4.4	326	2	F96522	hypothetical prote
9	100.5	4.3	514	2	T10559	hypothetical prote
10	100	4.3	1218	2	T30447	probable helicase
11	98.5	4.2	537	2	JC7127	frizzled protein 4
12	96	4.1	886	2	A57172	probable hormone r
13	93.5	4.0	254	2	T28170	hypothetical prote
14	93	4.0	630	2	JQ1894	hypothetical 74.1K
15	92.5	4.0	423	1	E71065	hypothetical prote
16	92	4.0	1014	2	T18759	hypothetical prote
17	91	3.9	285	2	E70473	phosphate transpor
18	91	3.9	288	2	S36956	cytochrome-c oxida
19	91	3.9	320	1	A05023	cytochrome c-type
20	91	3.9	322	2	T06955	probable cytochrom
21	90.5	3.9	470	2	B24670	probable a factor
22	90.5	3.9	2731	1	VFIHJH	genome polyprotein
23	90	3.9	455	1	GQHUT1	tumor necrosis fac
24	89	3.8	813	2	G86282	protein F10B6.32 [
25	88.5	3.8	343	2	E64882	probable membrane
26	88.5	3.8	343	2	A85751	hypothetical prote
27	88.5	3.8	517	2	S23209	mRNA maturase bI3
28	88.5	3.8	1164	2	T01871	RNA-directed DNA p
29	88	3.8	1107	2	T20578	hypothetical prote
30	88	3.8	1119	2	T20577	hypothetical prote
31	87.5	3.8	1311	2	C84528	hypothetical prote
32	87	3.7	308	2	T25508	hypothetical prote
33	87	3.7	364	2	T24657	hypothetical prote
34	87	3.7	394	2	T31891	hypothetical prote
35	87	3.7	490	2	T15875	hypothetical prote
36	87	3.7	535	2	T38244	hypothetical prote
37	87	3.7	1646	1	WMTMS2	186K protein - cuc
38	86.5	3.7	314	2	T21003	hypothetical prote
39	86.5	3.7	343	2	T33945	hypothetical prote
40	86.5	3.7	374	2	S07845	mRNA maturase bI3
41	86.5	3.7	507	2	T27627	hypothetical prote
42	86.5	3.7	592	2	D84431	probable endosomal
43	86.5	3.7	625	2	S60401	nuclear division p
44	86	3.7	237	2	C40784	hypothetical prote
45	86	3.7	402	2	S23774	triose phosphate/3

ALIGNMENTS

RESULT 1
T00250
MEGF2 protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C;Accession: T00250
R;Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998
A;Title: Identification of high-molecular-weight proteins with multiple EGF-like moti
A;Reference number: Z14126; MUID:98360089
A;Accession: T00250
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1364 <NAK>
A;Cross-references: EMBL:AB011536; NID:g3449297; PIDN:BAA32464.1; PID:g3449298
A;Experimental source: brain; clone HG1044
C;Genetics:
A;Gene: MEGF2
A;Map position: 3p21.2-p24.1
C;Superfamily: unassigned EGF-related proteins; EGF homology; laminin-type EGF-like h
F;1-28/Domain: EGF homology (fragment) <EGF>
F;32-66/Domain: EGF homology <EGF1>
F;124-169/Domain: laminin-type EGF-like homology <LEG>

Query Match 5.1%; Score 118; DB 2; Length 1364;
Best Local Similarity 22.1%; Pred. No. 0.058;
Matches 50; Conservative 33; Mismatches 93; Indels 50; Gaps 9;

```
Qy    207 MWIIR-----HRCTSH----------------ISP-----GSLEIL--------IITMIC 232
          :|  |      ||  ||                ||      | ||:|         :::
Db    546 VWTARDCELVHRNGSHARCRCSRTGTFGVLMDASPRERLEGDLELLAVFTHVVVAVSVAA 605

Qy    233 FVLTIAVYLYIKKLR-NVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFR 291
          ||| |: | :: |: || |            : | ::  |      :|:      |:|
Db    606 LVLTAAILLSLRSLKSNVRGIHANVAAALGVAELLFLLGIHRTHNQLVCTAVAILLHYFF 665

Qy    292 MASNLWLSVISYHTWKVLTSLNRVDPN-YRFLRYNAFVWSTAAIMTGSIYIVNQIWENDP 350
          :::  || |   | :::      ||    ||  |:|  |   |:: |    ::     :|
Db    666 LSTFAWLFVQGLHLYRMQVEPRNVDRGAMRF--YHALGWGVPAVLLGLAVGLDPEGYGNP 723

Qy    351 SKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTA 396
                  | :    | :| : :|| : :   |  || | |
Db    724 D----------FCWISVHEPLIWSF-AGPVVLVIVMNGTMFLLAA 757
```

RESULT 2
T15308
hypothetical protein B0286.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15308

```
R;Johnson, D.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid B0286.
A;Reference number: S69011
A;Accession: T15308
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1582 <JOH>
A;Cross-references: EMBL:U39848; NID:g1053217; PID:g1053218; PIDN:AAA80689.1; CESP:B0286
C;Genetics:
A;Gene: CESP:B0286.2
A;Introns: 54/2; 71/3; 126/3; 164/1; 216/1; 310/3; 347/2; 413/2; 479/2; 499/1; 574/2; 68

  Query Match           5.0%;  Score 116;  DB 2;  Length 1582;
  Best Local Similarity  20.6%;  Pred. No. 0.1;
  Matches   41;  Conservative   44;  Mismatches   80;  Indels   34;  Gaps   12;

Qy    228 ITMICFVLTIAVYLYIKKLRNVTGK-----CIVCCIVSRFIQCLIMILDHLNLLNGICSP 282
          |:::|  |:: |: : : |:||        |: | :::  :   :: :|     | |
Db    917 ISIVCLALSVCVFTFFRNLQNVRNSIHRNLCL-CLLIAELV--FVIGMDRTGNRTG-CGV 972

Qy    283 AGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPN-YRFLRYNAFVWSTAAIMTGSIYI 341
            |:| ::|  |: :  |  : :|  :   :||  |   |  | : | |::
Db    973 VAILLHYFFLSSFCWMLLEGYQLYMMLIQV--FEPNRTRIFLYYLFCYGTPAVVVAISAG 1030

Qy    342 VNQIWEN---DPSKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIY 398
          :   ||:   |   |           :    |::| ::: | : :   |: :| | |:
Db   1031 IK--WEDYGTDSYCW-----------IDTSTPTIWAFVA-PIIVIIAANI-IFLLIAL- 1073

Qy    399 IRKVKGGINKFTNEEEGRI 417
            ||   :      : |||
Db   1074 --KVVLSVQSRDRTKWGRI 1090


RESULT   3
C42523
A55R protein - vaccinia virus (strain Copenhagen)
C;Species: vaccinia virus
A;Note: host Homo sapiens (man)
C;Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 31-Oct-1997
C;Accession: C42523
R;Johnson, G.P.
submitted to GenBank, June 1990
A;Reference number: A33172
A;Accession: C42523
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-564 <JOH>
C;Superfamily: A55R protein; A55R protein carboxyl-terminal homology; A55R protein middl
F;7-110/Domain: POZ domain homology <POZ>
F;301-374/Domain: A55R protein middle region homology <MRH>
F;381-554/Domain: A55R protein carboxyl-terminal homology <CRH>

  Query Match           4.7%;  Score 109.5;  DB 2;  Length 564;
  Best Local Similarity  22.9%;  Pred. No. 0.12;
  Matches   94;  Conservative   49;  Mismatches  130;  Indels  137;  Gaps   24;

Qy     14 LIFTNNSNADIPGCNYYDTVDISYIERQNDSYLYDDI-EIPASLT--------GYYEFRQ 64
          ::|:||        | |: |::|::: |: : | | |  || |||           :| |
Db     48 ILFSNN----FIDSNEYE-VNLSHLDYQSVNDLIDYIYGIPLSLTNDNVKYILSTADFLQ 102

Qy     65 FGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLKEELARFKPYIYFTYMD 124
            |  :||  | ::   :|| : ||      |  :  | |      | | | |   :
Db    103 IG-SAITECENYILKNLCS-KNCIDFYIYADKY--NNK-----KIESASFNTILQ----- 148

Qy    125 LQARVPLTDMAIIRDEFFDCDEMIYISDFNYFLEEVSIQI------------FNKCGLIV 172
                : :| ||           :| |  ||  |:|            |    ||
Db    149 -------NILRLINDE-----------NFKYLTEESMIKILSDDMLNIKNEDFAPLILIK 190

Qy    173 WFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIITMIC 232
          | :  :   ||:|       ||          ||::   ||  |          |   |
Db    191 WLESTQQSCTVELLK-----CLRISLLSPQVIKSLY--SHRLVSS---------IYECIT 234

Qy    233 FVLTIA----------VYLYIKKLRNVTGKCIVCC----------IVSRFIQC--LIMI 269
          |:  ||           : |    : |   |  : |          | ||  :|   : :
Db    235 FLNNIAFLDESFPRYHSIELISIGISNSRDKISINCYNHKKNTWEMISSRRYRCSFAVAV 294

Qy    270 LDHLNLLNGICSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVD-PNYRFLRYNAFV 328
          ||::  : |    ||    :| :      ||:|:|    |:    | |  :: | |
Db    295 LDNIIYMMG-----GYDQSPYRSS-----KVIAYNT---CTNSWIYDIPELKYPRSNC-- 339

Qy    329 WSTAAIMTGS------IYIVNQIWENDPSKWNWLPLVGFIRCSVKDWHPS 372
                |       || :  | : | |          :  |:  | ||
Db    340 --------GGLADDEYIYCIGGIRDQDSS----------LTSSIDKWKPS 371


RESULT   4
T29673
hypothetical protein C50E3.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29673
R;Geisel, C.; Bradshaw, H.
submitted to the EMBL Data Library, July 1996
A;Description: The sequence of C. elegans cosmid C50E3.
A;Reference number: Z20662
A;Accession: T29673
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-323 <GEI>
A;Cross-references: EMBL:U64848; PIDN:AAB04882.1; GSPDB:GN00023; CESP:C50E3.7
A;Experimental source: strain Bristol N2; clone C50E3
C;Genetics:
A;Gene: CESP:C50E3.7
A;Map position: 5
A;Introns: 111/3; 243/3; 300/3; 316/3

  Query Match           4.5%;  Score 105.5;  DB 2;  Length 323;
  Best Local Similarity  19.9%;  Pred. No. 0.14;
  Matches   62;  Conservative   58;  Mismatches   94;  Indels   97;  Gaps   17;

Qy    132 TDMAIIRDEFFDCDEMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFWVTVDL----FM 187
          |:|  :      |::  |: :    |   | | |     :|      | ::| |    |:
Db     44 TNMKDLNCPRLKCEDEYYVKELVVALGNSSSQSFLSIAGLVHLAVNLFNISVQLSCVKFI 103

Qy    188 EKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVYLYIKKLR 247
          :|:          : | || :                :||:|   |||::: |:|:|  :
Db    104 DKR----------ALFKKSTF--------------FKILLIQ--CFVVSLRVFLHILTI- 136

Qy    248 NVTGKCIVCCI------VSRFI-QCLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLSV 300
            .      ||:  :      ||:|: || : :               | |:|| :     :|:
Db    137 -----CIIIYVSNTESSVSKFLSQCSLYV-------------DYCSNFFSLTVTFLMSL 177

Qy    301 ------ISYHTW--KVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIV-NQIWENDPS 351
                :|| ||  ::    | :        |:  : :  :|::  : |: :||
Db    178 NRCFCFVSY-TWNARIFDGKNVI--------YSVSIGAVISIVSAVLCIITSQI------ 222

Qy    352 KWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYI-------RKVKG 404
          : |:| : ||:     |  | :   :    |   | : |  | |:       :: :
Db    223 RRNFLTITGFV-----DMGPDIGFKV----LINRLFFIFPFGSIACYVILFFVIRKQNQQ 273

Qy    405 GINKFTNEEEG 415
            : | :|   |
Db    274 ALTKTSNRSRG 284


RESULT   5
T16663
hypothetical protein R03G5.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C;Accession: T16663
R;Connell, M.
```

```
submitted to the EMBL Data Library, March 1996
A;Description: The sequence of C. elegans cosmid R03G5.
A;Reference number: Z18556
A;Accession: T16663
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1209 <CON>
A;Cross-references: EMBL:U51994; NID:g1255290; PID:g1255293; PIDN:AAA96065.1; GSPDB:GN00
A;Experimental source: strain Bristol N2; clone R03G5
C;Genetics:
A;Gene: CESP:R03G5.3
A;Map position: X
A;Introns: 22/1; 103/1; 140/2; 194/1; 227/3; 276/1; 434/3; 615/3; 633/2; 649/3; 694/1; 7

  Query Match             4.5%;  Score 105;  DB 2;  Length 1209;
  Best Local Similarity   18.9%;  Pred. No. 0.68;
  Matches   81;  Conservative   67;  Mismatches  131;  Indels  150;  Gaps   24;

Qy      18 NNSNADIPG-----CNYYDTVDISYIERQNDSYLYDDIEIPASLTGYYEFRQ------FG 66
           | :|  :|       :   || | |       |: :: |:  |   ::: |       ||
Db     714 NVTNLGMPNLKSVIADVLPTVRIVY-----GLYIQENTELCLSEEEFWKLRNFQRSFLFG 768

Qy      67 DGSITPIEKHLRACVCSVRPCIRICCPAKNF-------------------LANGKCDDG 106
             |: ||              |   :|                      | |
Db     769 MKICCPVTKHY-------------CSQSSFQYFDMPLGCQILLSGIWLRDLTNTTYQSH 814

Qy     107 LKEELARFK----PYIYFTYMDLQARVPLTDMAI-----IRDEFFDCDE----MIYISDF 153
           | :  :  :    | : |       |  ||:::      || :|:|       : |
Db     815 LNQAFSTVEEVHGPIVVF-------RTDLTNLSFPNLHTIRSQFYDGQNDWQWQVLIIRG 867

Qy     154 NYFLEEVSIQIFNKCGLIVWFQDG-------------------KFWVTVDLFMEKQDY- 192
           |||| |    |:|  |  |:  |                     | :: :  |  |  |
Db     868 NYFLTSVE---FSK--LKNWYHFGREPQLTNSANYIWNYTPENCKLFLKIPFFFIKDTYS 922

Qy     193 -CLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEIL---IITMICFVLTIAVYLYIKKLRN 248
            |::| :    | |  ::    | :     |:      |:|:|  :||::  |::
Db     923 NCVFRESI-FFEYKLYFVFISFCATFQEQRSVNFFHFEIVTLI--LLTMSENLFVN---- 975

Qy     249 VTGKCIVCCIVSRFIQCLIMILDHLNLL--NGICSPAGYSSHFFRMASNLW--LSVISY- 303
                                :|::::   ||   | |  | |    : |:  :| : |.
Db     976 --------------------IDYIHIFDDNGFAIPFGLRSIFLNRDNALFNQISNVMYA 1014

Qy     304 HTWKVLTSLNRVDPNY-RFLRYNAFVWSTAAIMTGSIYIVNQIW--------ENDPS--- 351
           :   :| | :   | |   | |  :    | :::  :::|: :|        :|: :
Db    1015 NRLALLMSCSLYRPFYLTNLNYEKY----ALLVSCMVFVVSNVWNLSVMETLQNNFAYVL 1070

Qy     352 -KWNWLPLV 359
            : ||| |:
Db    1071 FQMNWLTLI 1079


RESULT   6
JQ1792
SalF17R protein - vaccinia virus
N;Alternate names: SalF21R 64.7K protein
C;Species: vaccinia virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000
C;Accession: JQ1792; C38550
R;Smith, G.L.; Chan, Y.S.; Howard, S.T.
J. Gen. Virol. 72, 1349-1376, 1991
A;Title: Nucleotide sequence of 42kbp of vaccinia virus strain WR from near the right ir
A;Reference number: JQ1767; MUID:91259063
A;Accession: JQ1792
A;Molecule type: DNA
A;Residues: 1-564 <SMI>
A;Cross-references: DDBJ:D11079; NID:g222717; PIDN:BAA01828.1; PID:g222743
A;Experimental source: strain WR
R;Howard, S.T.; Chan, Y.S.; Smith, G.L.
Virology 180, 633-647, 1991
A;Title: Vaccinia virus homologues of the shope fibroma virus inverted terminal repeat p
A;Reference number: A38550; MUID:91111982
A;Accession: C38550
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-564 <HOW>
A;Cross-references: GB:M58052
A;Experimental source: strain WR
C;Superfamily: A55R protein; A55R protein carboxyl-terminal homology; A55R protein mi
F;7-110/Domain: POZ domain homology <POZ>
F;301-374/Domain: A55R protein middle region homology <MRH>
F;381-554/Domain: A55R protein carboxyl-terminal homology <CRH>

  Query Match             4.5%;  Score 104.5;  DB 2;  Length 564;
  Best Local Similarity   22.4%;  Pred. No. 0.32;
  Matches   97;  Conservative   51;  Mismatches  136;  Indels  149;  Gaps   25;

Qy       8 FCTVLLLIFTNNSNAD---IPGC----------NYYDT----VDISYIERQNDSYLYDDI 50
           || : ::|    ||    : |          |: |:    |::|::: |: : | | |
Db      20 FCDISIVINDERINAHKLILSGASEYFSILFSNNFIDSNEYEVNLSHLDYQSVNDLIDYI 79

Qy      51 -EIPASLT--------GYYEFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANG 101
             || |||           :| | |  :|| | ::   :|| : ||     |  :  |
Db      80 YGIPLSLTNDNVKYILSTADFLQIG-SAITECENYILKNLCS-KNCIDFYIYADKY--NN 135

Qy     102 KCDDGLKEELARFKPYIYFTYMDLQARVPLTDMAIIRDEFFDCDEMIYISDFNYFLEEVS 161
           |    | | | |   :              : :| ||           :| |  ||
Db     136 K-----KIESASFNTILQ------------NILRLINDE----------NFKYLTEESM 167

Qy     162 IQI------------FNKCGLIVWFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWI 209
           |:|             |    || | :  :   ||:|       ||          ||::
Db     168 IKILSDDMLNIKNEDFAPLILIKWLESTQQSCTVELLR-----CLRISLLSPQVIKSLY- 221

Qy     210 IRHRCTSHISPGSLEILIITMICFVLTIA----------VYLYIKKLRNVTGKCIVCC- 257
                ||     |   |   | |:  ||           : |    : |   |  : |
Db     222 ------SH----QLVSSIYECITFLNNIAFLDESFPRYHSIELISIGISNSHDKISINCY 271

Qy     258 ---------IVSRFIQC--LIMILDHLNLLNGICSPAGYSSHFFRMASNLWLSVISYHTW 306
                  | ||  :|   : :||::  : |     ||    :| :       ||:|:|
Db     272 NHKKNTWEMISSRRYRCSFAVAVLDNIIYMMG-----GYDQSPYRSS-----KVIAYNT- 320

Qy     307 KVLTSLNRVD-PNYRFLRYNAFVWSTAAIMTGS------IYIVNQIWENDPSKWNWLPLV 359
             |:    | |  :: | |           |       || :  | : | |
Db     321 --CTNSWIYDIPELKYPRSNC----------GGLADDEYIYCIGGIRDQDSS-------- 360

Qy     360 GFIRCSVKDWHPS 372
             :  |:  | ||
Db     361 --LTSSIDKWKPS 371


RESULT   7
T14119
seven-pass transmembrane receptor protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14119
R;Hadjantonakis, A.K.; Formstone, C.J.; Little, P.F.R.
submitted to the EMBL Data Library, October 1997
A;Description: The Celsr family of novel evolutionarily conserved seven-pass transmem
A;Reference number: Z17881
A;Accession: T14119
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-3034 <HAD>
A;Cross-references: EMBL:AF031572; NID:g3800735; PID:g3800736; PIDN:AAC68836.1
C;Genetics:
A;Gene: Celsr1
A;Map position: 15
C;Keywords: transmembrane protein
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-3034/Product: seven-pass transmembrane receptor protein #status predicted <MAT>

  Query Match             4.4%;  Score 102;  DB 2;  Length 3034;
  Best Local Similarity   22.1%;  Pred. No. 3.5;
```

```
  Matches   38;  Conservative   31;  Mismatches   85;  Indels   18;  Gaps    6;

Qy     226 LIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGY 285
           |:  :: |||   |      | ::    |       | | : |:  :      :|:
Db    2495 LVALLVAFVLLSLVRTLRSNLHSIHKNLIAALF---FSQLIFMVGINQTENPFLCTVVAI 2551

Qy     286 SSHFFRMASNLWLSVISYHTWKVLTSLNRVDPN-YRFLRYNAFVWSTAAIMTGSIYIVNQ 344
             |:  | :  |  | : | :::|| :  :|    ||  |:  .|  .||:||    ::
Db    2552 LLHYVSMGTFAWTLVENLHVYRMLTEVRNIDTGPMRF--YHVVGWGIPAIVTGLAVGLDP 2609

Qy     345 IWENDPSKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTA 396
             :|  : || |   :          :| : :||   :   |  :| |:|
Db    2610 QGYGNPD-FCWLSLQDTL----------IWSF-AGPVGTVIIINTVIFVLSA 2649


RESULT   8
F96522
hypothetical protein F11A17.17 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: F96522
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719
A;Accession: F96522
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-326 <STO>
A;Cross-references: GB:AE005173; NID:g5733881; PIDN:AAD49769.1; GSPDB:GN00141
C;Genetics:
A;Gene: F11A17.17
A;Map position: 1


  Query Match             4.4%;  Score 101.5;  DB 2;  Length 326;
  Best Local Similarity   22.1%;  Pred. No. 0.31;
  Matches   31;  Conservative   30;  Mismatches   68;  Indels   11;  Gaps    5;

Qy     218 ISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLN 277
           |: |:  :  :     ||    |   |:||  : | :    :|  :    :|:   :
Db      22 INTGASSLSFVGSAFIVL---CYCLFKELRKFSFKLVFYLALSDMLCSFFLIVGDPS-KG 77

Qy     278 GICSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTG 337
            ||   ||::||| :|| || : |::   : :   :: |       :: :|| |: ::|
Db      78 FICYAQGYTTHFFCVASFLWTTTIAFTLHRTVVK-HKTDVEDLEAMFHLYVWGTSLVVT- 135

Qy     338 SIYIVNQIWENDP--SKWNW 355
             ::     |      | |
Db     136 ---VIRSFGNNHSHLGPWCW 152


RESULT   9
T10559
hypothetical protein F25E4.10 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999
C;Accession: T10559
R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16533
A;Accession: T10559
A;Molecule type: DNA
A;Residues: 1-514 <BEV>
A;Cross-references: EMBL:AL050399; GSPDB:GN00062; ATSP:F25E4.10
A;Experimental source: cultivar Columbia; BAC clone F25E4
C;Genetics:
A;Gene: ATSP:F25E4.10
A;Map position: 4
A;Introns: 436/3; 457/3; 479/3


  Query Match             4.3%;  Score 100.5;  DB 2;  Length 514;
  Best Local Similarity   19.5%;  Pred. No. 0.64;
  Matches   64;  Conservative   38;  Mismatches   93;  Indels  133;  Gaps   17;

Qy      92 CPAKNFLANGKCDDGLKEELARFKPYI--------YFTYMDLQARVPLTDMAIIRDEFFD 143
           |  :| : |||  :|: ||:   |||:         :|:: :   |  :      ||
Db     115 CATRNDVWNGKELEGVPEEIEDIKPYVVIYGNIIQHFSHKEHYLRFHVN--GFFRDGNKR 172

Qy     144 CDEMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFWVTVDLFMEKQDYCLYRHNFDSDF 203
           |    |          : :| | :|              :|      |: |  |   ::|
Db     173 CSACTY---------PIGLQSFYEC-------------IDC-----DFML--HQNCAEF 202

Qy     204 PKSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFI 263
           |:  | : |                 :  | ::| : :           || |  :|
Db     203 PRKKWHMLHN---------------DRLTLVTSVAEFFW----------CIACRRMSNGF 237

Qy     264 --QCLIMILDHLNLLNGICS---PAGYSSH------FFRMASNLWLSVISYHTWKVLTS- 311
             ||   :|| |      ||   |  : ||      : |          :  |: |||
Db     238 KYQCGYGVLDVL-----CCSVSEPFVHPSHPDHPLYYIRQVKYRQCYGCNGRTYNVLTCI 292

Qy     312 ------------------LNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWENDPSK 352
                            ::||| :   | |                        |   |
Db     293 ESCCGFVLCFCCATLPQVVMHRVDDHPLSLCYG--------------------EKASGK 331

Qy     353 WNWLPLVGFIRCSV--KDWHPSVWIYIS 378
           : |        | :  |: :|| | | |
Db     332 Y-W--------CDICEKELNPSTWFYTS 350


RESULT  10
T30447
probable helicase (EC 3.6.1.-) - Lymantria dispar nuclear polyhedrosis virus
C;Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000
C;Accession: T30447
R;Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; R
Virology 253, 17-34, 1999
A;Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantri
A;Reference number: Z20836; MUID:99124785
A;Accession: T30447
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1218 <KUZ>
A;Cross-references: EMBL:AF081810; NID:g3822234; PIDN:AAC70283.1; PID:g3822332
C;Superfamily: AcMNPV helicase
C;Keywords: hydrolase


  Query Match             4.3%;  Score 100;  DB 2;  Length 1218;
  Best Local Similarity   19.9%;  Pred. No. 1.9;
  Matches   58;  Conservative   44;  Mismatches   90;  Indels  100;  Gaps   14;

Qy      37 YIERQNDSYLYDDIEIPASLT----------GYYEFRQFGDG----------------SI 70
           |  :   :|| ||| :   :|                  ||||:                :|
Db     243 YSNKNFFNYLRDDINLVECVTAPRYQSLIRVNLKSLRQFGEAHSECKLDNFEALKITKNI 302

Qy      71 TP-------IEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLKEELARFKPYIYFTYM 123
           ||       |: |:::|: ::  |: :   :|  :|||     |::   :   :| | |:
Db     303 TPSSESLDEIKAHVKSCLAAIENCM-VSALSKLDIANGPV---LRDYFCK-SEFINFDYL 357

Qy     124 DLQARVPLTDMAIIRDEFFDCDEMIYISDFNYFLEEVSIQIFNK---------------- 167
             | |    ::::: ||  |    :|   :|| :   :|:
Db     358 -----VVLVWRLLVKNQEFDFAE----ADIRLYLELLCEALFSTPAGPAKVGAEQKTEKA 408

Qy     168 -----------CGLI--VWFQDGKFWVTVDLFMEKQ-----------DYCLYRHNFD--- 200
                      | |   :: :    |    ||||:             | || :|
Db     409 HEAALKRCEPYCNLTPKIFMRFCNHWT---LFMEEDALESIGCYFAIHYALYYNNTKRNP 465
```

```
Qy    201 --SDFPKSMWIIRHRCTSHISP-----GSLEILIITMICFVLTIAVYLYIKK 245
          |:  |  :       |  ::|     | |: :|    | |     |  :||
Db    466 TASERDKWNYTFESAVTCGLAPEVLARGFLKKIISANACLVFNGKHYTVVKK 517


RESULT  11
JC7127
frizzled protein 4 - human
C;Species: Homo sapiens (man)
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 31-Mar-2001
C;Accession: JC7127
R;Kirikoshi, H.; Sagara, N.; Koike, J.; Tanaka, K.; Sekihara, H.; Hirai, M.; Katoh, M.
Biochem. Biophys. Res. Commun. 264, 955-961, 1999
A;Title: Molecular cloning and characterization of human frizzled-4 on chromosome 11q14-
A;Reference number: JC7127; MUID:20012777; PMID:10544037
A;Accession: JC7127
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-537 <KIR>
A;Cross-references: GB:AB032417; DDBJ:AB032417; NID:g6277265; PID:g6277266
C;Genetics:
A;Gene: FZD4
A;Map position: 11 region q14 - q21
C;Superfamily: fruit fly frizzled protein


  Query Match             4.2%;  Score 98.5;  DB 2;  Length 537;
  Best Local Similarity   20.1%;  Pred. No. 0.99;
  Matches   85;  Conservative   60;  Mismatches  148;  Indels  129;  Gaps   22;

Qy     46 LYDDIEIPASLTGYYEFRQFGDGSITPIEKHLRACVCSV------------RPCIRICC 92
          |  | |:   || :    |:|  |       |:  :|||             ||  :|
Db     71 LQTDAEL--QLTTFTPLIQYGCSS------QLQFFLCSVYVPMCTEKINIPIGPCGGMCL 122

Qy     93 PAKNFLANGKCDDGLKE---------ELARFKPYIYFTYMDLQA----RVPLTDMAIIRD 139
            |      :|:  |||           ::| |     :| ::      |||     |:
Db    123 SVKR-----RCEPVLKEFGFAWPESLNCSKFPPQNDHNHMCMEGPGDEEVPLPHKTPIQP 177

Qy    140 EFFDCDEMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKF----------WVTV--DLFM 187
            :|  :   ||  |   : |:    |||    :  | :           |: |   |
Db    178 G-EECHSVGTNSD-QYIWVKRSLNCVLKCG----YDAGLYSRSAKEFTDIWMAVWASLCF 231

Qy    188 EKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVYLYIKKLR 247
          : :     ||              |  |     |: ::|   : :||   || :|
Db    232 ISTAFTVLTFLIDS-------------SRFSYPERPIIFLSMCYNIYSIA---YIVRL- 273

Qy    248 NVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLSVISYHTWK 307
            |:  : |    | :    :|    | |  |:      :|| |||::|  :::  ||
Db    274 -TVGRERISC---DFEEAAEPVLIQEGLKNTGCAIIFLLMYFFGMASSIWWVILTL-TWF 328

Qy    308 VLTSLN------RVDPNYRFLRYNAFVWSTAAI--------------MTGSIYIVNQIW 346
          :  |        :  :|    ::   |:  |:               :||  |: ||
Db    329 LAAGLKWGHEAIEMHSSY----FHIAAWAIPAVKTIVILIMRLVDADELTGLCYVGNQ-- 382

Qy    347 ENDPSKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGI 406
              |   | ||:          :: |:   :| ::   ||:|        |::  :
Db    383 -------NLDALTGFVVA-------PLFTYLVIGTLFIAAGLVALF--------KIRSNL 420

Qy    407 NK 408
            |
Db    421 QK 422


RESULT  12
A57172
probable hormone receptor EMR1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 11-Jan-2000
C;Accession: A57172
R;Baud, V.; Chissoe, S.L.; Viegas-Pequignot, E.; Diriong, S.; N'Guyen, V.C.; Roe, B.A.;
Genomics 26, 334-344, 1995
A;Title: EMR1, an unusual member in the family of hormone receptors with seven transmemb
A;Reference number: A57172; MUID:95324926
A;Accession: A57172
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-886 <BAU>
A;Cross-references: GB:X81479; NID:g784993; PIDN:CAA57232.1; PID:g784994
C;Genetics:
A;Gene: GDB:EMR1
A;Cross-references: GDB:378349; OMIM:600493
A;Map position: 19p13.3-19p13.3
C;Superfamily: unassigned EGF-related proteins; EGF homology
C;Keywords: transmembrane protein
F;136-170/Domain: EGF homology <EGF>


  Query Match             4.1%;  Score 96;  DB 2;  Length 886;
  Best Local Similarity   21.0%;  Pred. No. 2.9;
  Matches   48;  Conservative   40;  Mismatches   75;  Indels   66;  Gaps   11;

Qy    202 DFPKSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSR 261
          ||. |::||     ||:       :||:::| || || :|  : :||           :
Db    599 DF--SLYII-----SHVG------IIISLVCLVLAIATFLLCRSIRN----------HNT 635

Qy    262 FIQ---CLIMILDHLNLLNGI--------CSPAGYSSHFFRMASNLWLSVISYHTWKVLT 310
          ::    |: ::|     | ||         |:      |:  :|   |: | :   : ::
Db    636 YLHLHLCVCLLLAKTLFLAGIHKTDNKTGCAIIAGFLHYLFLACFFWMLVEAVILFLMVR 695

Qy    311 SLNRVD----PNYRFLRYNAF--------VWSTAAIMTGSIYIVNQIWENDPSKWNWLPL 358
          :|  |:     | : |   ||         |  :|::       : |: | |  :
Db    696 NLKVVNYFSSRNIKMLHICAFGYGLPMLVVVISASVQPQGYGMHNRCWLNTET------- 748

Qy    359 VGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGIN 407
           |||           | :: ||   :   |  :   |   :|:    :|
Db    749 -GFI----------WSFL-GPVCTVIVINSLLLTWTLWILRQRLSSVN 784


RESULT  13
T28170
hypothetical protein ORF9 - Melanoplus sanguinipes entomopoxvirus (strain Tuscon)
C;Species: Melanoplus sanguinipes entomopoxvirus
A;Variety: strain Tuscon
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C;Accession: T28170
R;Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
J. Virol. 73, 533-552, 1999
A;Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A;Reference number: Z20484; MUID:99102612
A;Accession: T28170
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-254 <AFO>
A;Cross-references: EMBL:AF063866; NID:g4049647; PIDN:AAC97858.1; PID:g4049898
A;Experimental source: strain Tuscon
C;Genetics:
A;Note: MSV009


  Query Match             4.0%;  Score 93.5;  DB 2;  Length 254;
  Best Local Similarity   22.4%;  Pred. No. 1.2;
  Matches   57;  Conservative   40;  Mismatches   78;  Indels   79;  Gaps   13;

Qy     75 KHLRACVCSVR-----------PCIRICCPAKNFLANGKCDDGLKEELARFKPYIYFTYM 123
          ||:|   || :            |: ||| :|
Db     53 KHVRKLQCSGKIPNNLEKLINLECLTICCASK-------------------------- 84

Qy    124 DLQARVPLTDMAIIRDEFFDCDEMIYISDFN-----YFLEEVSIQIFNK------CGL-- 170
            :::::|::   |:   ||:|     |:||      | ||:: | | ||      | |
Db     85 -IESKLPISLKKIV---FFNCK----ITDFKFLENLYNLEKLFI-INNKKSNISDCKLPI 135

Qy    171 -IVWFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIIT 229
           |: |   |  :|   |:|     ||         | ::|  ::  |:||  :| | :
Db    136 SIIEFCCNKCEITDFKFLED----LYN-------LKKLYISGNK-NSNISHCNLPIKLTE 183
```

```
Qy    230 MICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSS-H 288
          : |    |  : ::| | |:    |     |    | :       |:  :|     :  :
Db    184 LECSYCNILDFKFLKNLYNLKKLFITGNNNSNISNCEL----PFKLIKFVCINCNITDFN 239

Qy    289 FFRMASNLWLSVIS 302
          | : ::|| :  ||
Db    240 FLKSSNNLRIITIS 253


RESULT  14
JQ1894
hypothetical 74.1K protein - fowlpox virus
C;Species: fowlpox virus
C;Date: 21-Nov-1993 #sequence_revision 03-May-1994 #text_change 08-Oct-1999
C;Accession: JQ1894; S27934
R;Ogawa, R.; Calvert, J.G.; Yanagida, N.; Nazerian, K.
J. Gen. Virol. 74, 55-64, 1993
A;Title: Insertional inactivation of a fowlpox virus homologue of the vaccinia virus F12
A;Reference number: JQ1894; MUID:93139784
A;Accession: JQ1894
A;Molecule type: DNA
A;Residues: 1-630 <OGA>
A;Cross-references: GB:M88588; NID:g333522; PIDN:AAA47187.1; PID:g333524
C;Comment: This protein may be involved in the envelopment or release of infectious extr


  Query Match             4.0%;  Score 93;  DB 2;  Length 630;
  Best Local Similarity   22.0%;  Pred. No. 3.6;
  Matches   64;  Conservative   32;  Mismatches   89;  Indels  106;  Gaps   15;

Qy     13 LLIFT---NNSNADIPGCNYY-----DTVDISYIERQNDSYLYDDIE----------IP 53
          :|||:    |:|      |||       |::       | |  |||:            |
Db    400 ILIFSIRVKNNNG-----NYYYPKHSSAVNLMLSSIYTDYYAIDDIDKDSKKLVFNSIFP 454

Qy     54 ASLTGYYEFRQFGDG----SITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLKE 109
          : |||      :|    : || | :|  |:| |     |    || :             ||
Db    455 LIMEGYYP-----EGKPYYTKTPKEGYLSICLCDV----EISNDIKNPILY------CKE 499

Qy    110 ELARFKPYIYFTYMDLQARVPLTDMAIIRDEFFDCDEMIYISDFNYFLEEVSIQIFNKCG 169
          :  |    || :|:   | |    |   |  :    | : |        :|   ||
Db    500 NKSARKFTGVFTSVDIDTAVKLRGYKIKILECIEWPNKIKLFD--------NICYLNK-- 549

Qy    170 LIVWFQDGKFWVTVDLFMEKQDYC-------------LYRHNFDSDFPKSMWIIRHRCTS 216
                      ||:| |||               | : |   |      :    | :
Db    550 ---------------LFIEHQDYTHDEKSLQGYLFSYLLKGNVTEDV-----LAMKSCRN 589

Qy    217 HISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIV--SRFIQC 265
          ::|          :| |:::           || | | : | :   |  ::|
Db    590 NLS----------IISFIISYC--------RNYTYKLLECPVYESSNIVKC 622


RESULT  15
E71065
hypothetical protein PH1219 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: E71065
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137
A;Accession: E71065
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-423 <KAW>
A;Cross-references: GB:AP000005; NID:g3236132;, PIDN:BAA30319.1; PID:g3257636
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1219
C;Superfamily: Aquifex aeolicus glucose-1-phosphate thymidylyltransferase


  Query Match             4.0%;  Score 92.5;  DB 1;  Length 423;
  Best Local Similarity   21.4%;  Pred. No. 2.5;
  Matches   63;  Conservative   48;  Mismatches   98;  Indels   85;  Gaps   16;

Qy    122 YMDLQ--ARVPLTDMAIIR-----DEFFDCDEMIYISDFNYF------LEEVSI----QI 164
          |:|::   :| :  | ::       |:   |   :| | : |      ::|  :     :
Db    126 YVDVEEATKVKIEGMRVVEIGKKLKEWNAVDTGFFILDDSIFEITTRLVQEKEVVELRDV 185

Qy    165 FNKCGLIVWFQDGKFWVTVDL--FMEKQDYCLYRHN-------FDSDF---PKSMWIIRH 212
            :  | | | || ||: ||    ::|    :  |:       | | :     | ||  |
Db    186 VKEAKLKVTFVDGLFWMDVDTPEELKKAKKLIVAHSVKGVGDGFISRYLNRKVSTWISSH 245

Qy    213 RCTSHISPGSLEILIIT------MICFV-LTIAVYLY------------IKKLRNVTGKC 253
              |::|  : ::  |       :: || : :|  ||            | : |  |
Db    246 -LVDHVTPNQMTLITFTFGIISALVNFVNIPLAGILYQISSILDGVDGEIARARMQT--- 301

Qy    254 IVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLSV----------ISY 303
            |:|      |||    |: : : |     :  :|  || ||          :||
Db    302 ------SKFGGYFDSILDRYVDLSFLLTLA-----YVSLADPLWWSVATLAIFSSAMVSY 350

Qy    304 HT----------WKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIW 346
          |           :||: :| :: |  |  |     :  |      :::|:  ||
Db    351 STERFRGAYCADAYKVIPALRKI-PGKRDERIFTIMIFTLLGWIKALFILLAIW 403


Search completed: October 30, 2001, 15:21:50
Job time: 716 sec
```

THIS PAGE BLANK (USPTO)

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 15:20:44 ; Search time 56.92 Seconds
(without alignments)
256.375 Million cell updates/sec

Title: US-09-619-049-855
Perfect score: 2324
Sequence: 1 MRLPWVIFCTVLLLIFTNNS.....NKFTNEEEGRINCINFDSQT 426

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	109.5	4.7	564	1	VA55_VACCC	P21073 vaccinia vi
2	104.5	4.5	564	1	VA55_VACCV	P24768 vaccinia vi
3	96	4.1	886	1	EMR1_HUMAN	Q14246 homo sapien
4	93	4.0	630	1	VF12_FOWPV	P36317 fowlpox vir
5	93	4.0	931	1	EMR1_MOUSE	Q61549 mus musculu
6	91	3.9	320	1	CCSA_MARPO	P12214 marchantia
7	91	3.9	322	1	CCSA_CYAPA	P48257 cyanophora
8	90.5	3.9	470	1	STE3_YEAST	P06783 saccharomyc
9	90.5	3.9	2731	1	RRPB_CVMJH	P29982 murine coro
10	90	3.9	455	1	TNR1_HUMAN	P19438 homo sapien
11	88.5	3.8	343	1	YNAI_ECOLI	P77253 escherichia
12	88	3.8	553	1	GLP2_HUMAN	O95838 homo sapien
13	87	3.7	490	1	YYO1_CAEEL	Q18904 caenorhabdi
14	87	3.7	535	1	YDW6_SCHPO	O13912 schizosacch
15	86.5	3.7	471	1	5H2A_MACMU	P50128 macaca mula
16	86.5	3.7	625	1	TRF5_YEAST	P48561 saccharomyc
17	86.5	3.7	1648	1	RRPO_CGMVS	P19523 cucumber gr
18	86	3.7	402	1	CPTR_PEA	P21727 pisum sativ
19	86	3.7	452	1	YCDT_ECOLI	P75908 escherichia
20	85.5	3.7	410	1	YJIO_ECOLI	P39386 escherichia
21	85.5	3.7	477	1	CB33_YEAST	P35203 saccharomyc
22	85.5	3.7	517	1	MBI3_YEAST	Q9zzw7 saccharomyc
23	85.5	3.7	591	1	YKT4_CAEEL	P34315 caenorhabdi
24	85.5	3.7	2733	1	RRPB_CVMA5	P16342 murine coro
25	85	3.7	772	1	LP1G_DROME	P11997 drosophila
26	85	3.7	4488	1	RRPA_CVMJH	P19751 murine coro
27	84.5	3.6	270	1	Y304_MYCPN	P75355 mycoplasma
28	84.5	3.6	830	1	FAR1_YEAST	P21268 saccharomyc
29	84	3.6	381	1	VQ3L_CAPVK	Q86917 capripoxvir
30	84	3.6	405	1	YRKO_BACSU	P54442 bacillus su
31	83.5	3.6	359	1	HH2R_CAVPO	P47747 cavia porce
32	83.5	3.6	423	1	GRFR_PIG	P34999 sus scrofa
33	83.5	3.6	713	1	PRML_BRARE	Q9w735 brachydanio
34	83	3.6	1184	1	CHSE_EMENI	Q00744 emericella
35	82.5	3.5	423	1	GRFR_HUMAN	Q02643 homo sapien
36	82.5	3.5	464	1	GRFR_RAT	Q02644 rattus norv
37	82.5	3.5	471	1	5H2A_HUMAN	P28223 homo sapien
38	82	3.5	395	1	DIHR_MANSE	P35464 manduca sex
39	82	3.5	678	1	NTP2_SFVKA	Q9q927 shope fibro
40	81.5	3.5	346	1	YJ9S_YEAST	P47181 saccharomyc
41	81.5	3.5	352	1	OPSB_ORYLA	P87365 oryzias lat
42	81.5	3.5	471	1	5H2A_CRIGR	P18599 cricetulus
43	81.5	3.5	531	1	ACM5_RAT	P08911 rattus norv
44	81.5	3.5	1294	1	YA3B_SCHPO	Q09716 schizosacch
45	81.5	3.5	1541	1	MRP2_RAT	Q63120 rattus norv

ALIGNMENTS

```
RESULT   1
VA55_VACCC
ID   VA55_VACCC     STANDARD;      PRT;   564 AA.
AC   P21073;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROTEIN A55.
GN   A55R.
OS   Vaccinia virus (strain Copenhagen).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
OX   NCBI_TaxID=10249;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91021027; PubMed=2219722;
RA   Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA   Paoletti E.;
RT   "The complete DNA sequence of vaccinia virus.";
RL   Virology 179:247-266(1990).
RN   [2]
RP   COMPLETE GENOME.
RA   Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA   Paoletti E.;
RL   Virology 179:517-563(1990).
CC   -!- SIMILARITY: CONTAINS 6 KELCH REPEATS.
CC   -!- SIMILARITY: CONTAINS 1 BTB/POZ DOMAIN.
CC   -!- SIMILARITY: BELONGS THE POXVIRUSES KELCH FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M35027; AAA48190.1; -.
DR   PIR; C42523; C42523.
DR   InterPro; IPR000210; -.
DR   InterPro; IPR001798; -.
DR   Pfam; PF00651; BTB; 1.
DR   Pfam; PF01344; Kelch; 5.
DR   PRINTS; PR00501; KELCHREPEAT.
DR   PROSITE; PS50097; BTB; 1.
KW   Early protein; Repeat.
FT   DOMAIN       21     88       BTB.
FT   REPEAT      252    297       KELCH 1.
FT   REPEAT      298    346       KELCH 2.
FT   REPEAT      347    395       KELCH 3.
FT   REPEAT      397    441       KELCH 4.
FT   REPEAT      442    492       KELCH 5.
FT   REPEAT      494    539       KELCH 6.
SQ   SEQUENCE   564 AA;  64718 MW;  DBE8ECB532DA9F71 CRC64;
```

```
  Query Match          4.7%;  Score 109.5;  DB 1;  Length 564;
  Best Local Similarity   22.9%;  Pred. No. 0.088;
  Matches   94;  Conservative   49;  Mismatches  130;  Indels  137;  Gaps   24;

Qy      14 LIFTNNSNADIPGCNYYDTVDISYIERQNDSYLYDDI-EIPASLT--------GYYEFRQ 64
           ::|:||        | |: |::|::: |: : | | |  || |||           :| |
Db      48 ILFSNN----FIDSNEYE-VNLSHLDYQSVNDLIDYIYGIPLSLTNDNVKYILSTADFLQ 102

Qy      65 FGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLKEELARFKPYIYFTYMD 124
             |  :||  | ::   :|| : ||      |  :  | |      | | | |   :
Db     103 IG-SAITECENYILKNLCS-KNCIDFYIYADKY--NNK-----KIESASFNTILQ----- 148

Qy     125 LQARVPLTDMAIIRDEFFDCDEMIYISDFNYFLEEVSIQI------------FNKCGLIV 172
                  : :| ||           :| |  ||  |:|             |    ||
Db     149 -------NILRLINDE-----------NFKYLTEESMIKILSDDMLNIKNEDFAPLILIK 190

Qy     173 WFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIITMIC 232
           | :  :   ||:|       ||         ||::   ||  |          |   |
Db     191 WLESTQQSCTVELLK-----CLRISLLSPQVIKSLY--SHRLVSS---------IYECIT 234

Qy     233 FVLTIA-----------VYLYIKKLRNVTGKCIVCC----------IVSRFIQC--LIMI 269
           |:  ||           : |    : |   |  : |          | ||  :|   : :
Db     235 FLNNIAFLDESFPRYHSIELISIGISNSRDKISINCYNHKKNTWEMISSRRYRCSFAVAV 294

Qy     270 LDHLNLLNGICSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVD-PNYRFLRYNAFV 328
           ||::  : |     ||    :| :      ||:|:|    |:    | |  :: | |
Db     295 LDNIIYMMG-----GYDQSPYRSS-----KVIAYNT---CTNSWIYDIPELKYPRSNC-- 339

Qy     329 WSTAAIMTGS------IYIVNQIWENDPSKWNWLPLVGFIRCSVKDWHPS 372
                 |      || :  | : | |          :  |:  | ||
Db     340 --------GGLADDEYIYCIGGIRDQDSS----------LTSSIDKWKPS 371


RESULT   2
VA55_VACCV
ID   VA55_VACCV     STANDARD;      PRT;   564 AA.
AC   P24768;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROTEIN A55.
GN   A55R OR SALF17R OR SALF21R.
OS   Vaccinia virus (strain WR).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
OX   NCBI_TaxID=10254;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91259063; PubMed=2045793;
RA   Smith G.L., Chan Y.S., Howard S.T.;
RT   "Nucleotide sequence of 42 kbp of vaccinia virus strain WR from near
RT   the right inverted terminal repeat.";
RL   J. Gen. Virol. 72:1349-1376(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91111982; PubMed=1846491;
RA   Howard S.T., Chan Y.S., Smith G.L.;
RT   "Vaccinia virus homologues of the Shope fibroma virus inverted
RT   terminal repeat proteins and a discontinuous ORF related to the tumor
RT   necrosis factor receptor family.";
RL   Virology 180:633-647(1991).
CC   -!- SIMILARITY: CONTAINS 6 KELCH REPEATS.
CC   -!- SIMILARITY: CONTAINS 1 BTB/POZ DOMAIN.
CC   -!- SIMILARITY: BELONGS THE POXVIRUSES KELCH FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D11079; BAA01828.1; -.
DR   EMBL; M58054; AAA48340.1; -.
DR   PIR; C38550; C38550.
DR   PIR; JQ1792; JQ1792.
DR   InterPro; IPR000210; -.
DR   InterPro; IPR001798; -.
DR   Pfam; PF00651; BTB; 1.
DR   Pfam; PF01344; Kelch; 5.
DR   PRINTS; PR00501; KELCHREPEAT.
DR   PROSITE; PS50097; BTB; 1.
KW   Early protein; Repeat.
FT   DOMAIN       21     88       BTB.
FT   REPEAT      252    297       KELCH 1.
FT   REPEAT      298    346       KELCH 2.
FT   REPEAT      347    395       KELCH 3.
FT   REPEAT      397    441       KELCH 4.
FT   REPEAT      442    492       KELCH 5.
FT   REPEAT      494    539       KELCH 6.
SQ   SEQUENCE   564 AA;  64690 MW;  FDAC2001F6EF7507 CRC64;


  Query Match          4.5%;  Score 104.5;  DB 1;  Length 564;
  Best Local Similarity   22.4%;  Pred. No. 0.23;
  Matches   97;  Conservative   51;  Mismatches  136;  Indels  149;  Gaps   25;

Qy       8 FCTVLLLIFTNNSNAD---IPGC----------NYYDT----VDISYIERQNDSYLYDDI 50
           || : ::|     ||    : |           |: |:    |::|::: |: : | | |
Db      20 FCDISIVINDERINAHKLILSGASEYFSILFSNNFIDSNEYEVNLSHLDYQSVNDLIDYI 79

Qy      51 -EIPASLT--------GYYEFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANG 101
            || |||           :| | |  :||  | ::   :|| : ||     |  :  |
Db      80 YGIPLSLTNDNVKYILSTADFLQIG-SAITECENYILKNLCS-KNCIDFYIYADKY--NN 135

Qy     102 KCDDGLKEELARFKPYIYFTYMDLQARVPLTDMAIIRDEFFDCDEMIYISDFNYFLEEVS 161
           |     | | | |   :          . : :| ||          :| |  ||
Db     136 K-----KIESASFNTILQ-----------NILRLINDE----------NFKYLTEESM 167

Qy     162 IQI-----------FNKCGLIVWFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWI 209
           |:|           |    || | :  :   ||:|       ||          ||::
Db     168 IKILSDDMLNIKNEDFAPLILIKWLESTQQSCTVELLR-----CLRISLLSPQVIKSLY- 221

Qy     210 IRHRCTSHISPGSLEILIITMICFVLTIA----------VYLYIKKLRNVTGKCIVCC- 257
                ||     |   |   | |:  ||           : |    : |   |   : |
Db     222 ------SH----QLVSSIYECITFLNNIAFLDESFPRYHSIELISIGISNSHDKISINCY 271

Qy     258 ---------IVSRFIQC--LIMILDHLNLLNGICSPAGYSSHFFRMASNLWLSVISYHTW 306
                   | ||  :|   : :||::  : |    ||    :| :      ||:|:|
Db     272 NHKKNTWEMISSRRYRCSFAVAVLDNIIYMMG-----GYDQSPYRSS-----KVIAYNT- 320

Qy     307 KVLTSLNRVD-PNYRFLRYNAFVWSTAAIMTGS------IYIVNQIWENDPSKWNWLPLV 359
             |:    | |  :: | |           |       || :  | : | |
Db     321 --CTNSWIYDIPELKYPRSNC----------GGLADDEYIYCIGGIRDQDSS-------- 360

Qy     360 GFIRCSVKDWHPS 372
             :  |:  | ||
Db     361 --LTSSIDKWKPS 371


RESULT   3
EMR1_HUMAN
ID   EMR1_HUMAN     STANDARD;      PRT;   886 AA.
AC   Q14246;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CELL SURFACE GLYCOPROTEIN EMR1 PRECURSOR (EMR1 HORMONE RECEPTOR).
GN   EMR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE=95324926; PubMed=7601460;
RA   Baud V., Chissoe S.L., Viegas-Pequignot E., Diriong S., N'Guyen V.C.,
RA   Roe B.A., Lipinski M.;
RT   "EMR1, an unusual member in the family of hormone receptors with
RT   seven transmembrane segments.";
RL   Genomics 26:334-344(1995).
CC   -!- FUNCTION: PROBABLY INVOLVED IN CELLULAR RESPONSE TO A HORMONE OR
CC       AN INTERACTION WITH A PROTEIN LIGAND.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: WIDE EXPRESSION; INCREASED LEVELS IN
CC       PERIPHERAL BLOOD MONONUCLEAR CELLS.
CC   -!- PTM: N- AND O-GLYCOSYLATED; (POSSIBLE).
CC   -!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X81479; CAA57232.1; -.
DR   HSSP; P00736; 1APQ.
DR   MIM; 600493; -.
DR   InterPro; IPR000152; -.
DR   InterPro; IPR000203; -.
DR   InterPro; IPR000561; -.
DR   InterPro; IPR000832; -.
DR   InterPro; IPR001740; -.
DR   InterPro; IPR001881; -.
DR   Pfam; PF00002; 7tm_2; 1.
DR   Pfam; PF00008; EGF; 5.
DR   Pfam; PF01825; GPS; 1.
DR   PRINTS; PR01128; EMR1HORMONER.
DR   PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 6.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01187; EGF_CA; 5.
KW   G-protein coupled receptor; Transmembrane; Receptor; Glycoprotein;
KW   EGF-like domain; Repeat; Signal.
FT   SIGNAL       1     17       POTENTIAL.
FT   CHAIN       18    886       CELL SURFACE GLYCOPROTEIN EMR1.
FT   DOMAIN      18    599       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   600    627       POTENTIAL.
FT   DOMAIN     628    634       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   635    656       POTENTIAL.
FT   DOMAIN     657    666       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   667    690       POTENTIAL.
FT   DOMAIN     691    709       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   710    731       POTENTIAL.
FT   DOMAIN     732    747       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   748    776       POTENTIAL.
FT   DOMAIN     777    794       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   795    814       POTENTIAL.
FT   DOMAIN     815    829       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   830    852       POTENTIAL.
FT   DOMAIN     853    886       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      31     79       EGF-LIKE 1.
FT   DOMAIN      80    131       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     132    171       EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     172    213       EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     221    267       EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     268    316       EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     317    599       SER/THR-RICH.
FT   DISULFID    35     47       BY SIMILARITY.
FT   DISULFID    41     56       BY SIMILARITY.
FT   DISULFID    58     78       BY SIMILARITY.
FT   DISULFID    84     97       BY SIMILARITY.
FT   DISULFID    91    106       BY SIMILARITY.
FT   DISULFID   108    130       BY SIMILARITY.
FT   DISULFID   136    148       BY SIMILARITY.
FT   DISULFID   142    157       BY SIMILARITY.
FT   DISULFID   159    170       BY SIMILARITY.
FT   DISULFID   176    188       BY SIMILARITY.
FT   DISULFID   182    197       BY SIMILARITY.
FT   DISULFID   199    212       BY SIMILARITY.
FT   DISULFID   225    235       BY SIMILARITY.
FT   DISULFID   229    244       BY SIMILARITY.
FT   DISULFID   246    266       BY SIMILARITY.
FT   DISULFID   272    285       BY SIMILARITY.
FT   DISULFID   279    294       BY SIMILARITY.
FT   DISULFID   296    315       BY SIMILARITY.
FT   CARBOHYD    94     94       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    99     99       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   127    127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   167    167       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   189    189       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   194    194       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   232    232       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   258    258       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   312    312       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   366    366       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   375    375       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   448    448       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   661    661       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   886 AA;  97680 MW;  7456CA56FB624D99 CRC64;


  Query Match             4.1%;  Score 96;  DB 1;  Length 886;
  Best Local Similarity   21.0%;  Pred. No. 1.9;
  Matches   48;  Conservative   40;  Mismatches   75;  Indels   66;  Gaps   11;

Qy    202 DFPKSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSR 261
          ||  |::||     ||:       :||:::| || || :|  : :||           :
Db    599 DF--SLYII-----SHVG------IIISLVCLVLAIATFLLCRSIRN----------HNT 635

Qy    262 FIQ---CLIMILDHLNLLNGI--------CSPAGYSSHFFRMASNLWLSVISYHTWKVLT 310
          ::    |: ::|    | ||        |:      |:  :|   |: | :   : ::
Db    636 YLHLHLCVCLLLAKTLFLAGIHKTDNKTGCAIIAGFLHYLFLACFFWMLVEAVILFLMVR 695

Qy    311 SLNRVD----PNYRFLRYNAF--------VWSTAAIMTGSIYIVNQIWENDPSKWNWLPL 358
          :|  |:     | : |   ||        |  :|::      : |: | |  :
Db    696 NLKVVNYFSSRNIKMLHICAFGYGLPMLVVVISASVQPQGYGMHNRCWLNTET------- 748

Qy    359 VGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGIN 407
           |||           | :: ||   :   |  :   |   :|:    :|
Db    749 -GFI----------WSFL-GPVCTVIVINSLLLTWTLWILRQRLSSVN 784


RESULT  4
VF12_FOWPV
ID   VF12_FOWPV     STANDARD;      PRT;   630 AA.
AC   P36317;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   PROTEIN F12 HOMOLOG.
OS   Fowlpox virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139784; PubMed=8380837;
RA   Ogawa R., Calvert J.G., Yanagida N., Nazerian K.;
RT   "Insertional inactivation of a fowlpox virus homologue of the
RT   vaccinia virus F12L gene inhibits the release of enveloped virions.";
RL   J. Gen. Virol. 74:55-64(1993).
CC   -!- FUNCTION: IS INVOLVED IN THE ENVELOPMENT OR RELEASE OF INFECTIOUS
CC       EXTRACELLULAR VIRIONS.
CC   -!- SIMILARITY: SOME, TO PROTEIN F12 OF VACCINIA VIRUS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M88588; AAA47187.1; -.
DR   EMBL; M88587; AAA43820.1; -.
DR   PIR; JQ1894; JQ1894.
DR   PIR; S27934; S27934.
KW   Transmembrane; Glycoprotein.
FT   CARBOHYD     66     66       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     75     75       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    360    360       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    393    393       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    500    500       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    576    576       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    590    590       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   630 AA;  74098 MW;  CD8E1F102032020F CRC64;


  Query Match             4.0%;  Score 93;  DB 1;  Length 630;
  Best Local Similarity   22.0%;  Pred. No. 2.3;
  Matches   64;  Conservative   32;  Mismatches   89;  Indels  106;  Gaps   15;

Qy      13 LLIFT---NNSNADIPGCNYY-----DTVDISYIERQNDSYLYDDIE-----------IP 53
           :|||:    |:|      |||       |::     | |  |||:           |
Db     400 ILIFSIRVKNNNG-----NYYYPKHSSAVNLMLSSIYTDYYAIDDIDKDSKKLVFNSIFP 454

Qy      54 ASLTGYYEFRQFGDG----SITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLKE 109
             : |||      :|    : || | :|  |:| |     |    || :          ||
Db     455 LIMEGYYP-----EGKPYYTKTPKEGYLSICLCDV----EISNDIKNPILY------CKE 499

Qy     110 ELARFKPYIYFTYMDLQARVPLTDMAIIRDEFFDCDEMIYISDFNYFLEEVSIQIFNKCG 169
             :  |    || :|:   | |    |   |  :    | : |        :|   ||
Db     500 NKSARKFTGVFTSVDIDTAVKLRGYKIKILECIEWPNKIKLFD--------NICYLNK-- 549

Qy     170 LIVWFQDGKFWVTVDLFMEKQDYC-------------LYRHNFDSDFPKSMWIIRHRCTS 216
                      ||:| |||                | : |   |      :    | :
Db     550 --------------LFIEHQDYTHDEKSLQGYLFSYLLKGNVTEDV-----LAMKSCRN 589

Qy     217 HISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIV--SRFIQC 265
           ::|          :| |:::          || | | : | :   |  ::|
Db     590 NLS----------IISFIISYC--------RNYTYKLLECPVYESSNIVKC 622


RESULT   5
EMR1_MOUSE
ID   EMR1_MOUSE     STANDARD;      PRT;   931 AA.
AC   Q61549;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CELL SURFACE GLYCOPROTEIN EMR1 PRECURSOR (EMR1 HORMONE RECEPTOR)
DE   (CELL SURFACE GLYCOPROTEIN F4/80).
GN   EMR1 OR GPF480.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=Peritoneal cavity;
RX   MEDLINE=96132946; PubMed=8550607;
RA   McKnight A.J., Macfarlane A.J., Dri P., Turley L., Willis A.C.,
RA   Gordon S.;
RT   "Molecular cloning of F4/80, a murine macrophage-restricted cell
RT   surface glycoprotein with homology to the G-protein-linked
RT   transmembrane 7 hormone receptor family.";
RL   J. Biol. Chem. 271:486-489(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97312684; PubMed=9169125;
RA   Lin H.H., Stubbs L.J., Mucenski M.L.;
RT   "Identification and characterization of a seven transmembrane hormone
RT   receptor using differential display.";
RL   Genomics 41:301-308(1997).
CC   -!- FUNCTION: PROBABLY INVOLVED IN CELL ADHESION WITHIN TISSUES
CC       AND RECEPTOR SIGNALLING.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: IN MACROPHAGES; BUT ABSENT FROM THOSE WHICH
CC       ARE LOCALIZED WITHIN T-CELL AREAS OF LYMPH NODES AND SPLEEN.
CC       LOW LEVEL OF EXPRESSION ON BLOOD MONOCYTES.
CC   -!- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X93328; CAA63720.1; -.
DR   EMBL; U66888; AAC53184.1; -.
DR   HSSP; P07204; 1FGD.
DR   GCRDb; GCR_1309; -.
DR   MGD; MGI:106912; Emr1.
DR   InterPro; IPR000152; -.
DR   InterPro; IPR000203; -.
DR   InterPro; IPR000561; -.
DR   InterPro; IPR000832; -.
DR   InterPro; IPR001740; -.
DR   InterPro; IPR001881; -.
DR   Pfam; PF00002; 7tm_2; 1.
DR   Pfam; PF00008; EGF; 7.
DR   Pfam; PF01825; GPS; 1.
DR   PRINTS; PR01128; EMR1HORMONER.
DR   PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 6.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01187; EGF_CA; 5.
KW   G-protein coupled receptor; Transmembrane; Receptor; Glycoprotein;
KW   EGF-like domain; Repeat; Signal.
FT   SIGNAL       1     27       POTENTIAL.
FT   CHAIN       28    931       CELL SURFACE GLYCOPROTEIN EMR1.
FT   DOMAIN      28    644       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   645    672       POTENTIAL.
FT   DOMAIN     673    679       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   680    701       POTENTIAL.
FT   DOMAIN     702    711       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   712    735       POTENTIAL.
FT   DOMAIN     736    754       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   755    776       POTENTIAL.
FT   DOMAIN     777    792       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   793    821       POTENTIAL.
FT   DOMAIN     822    839       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   840    859       POTENTIAL.
FT   DOMAIN     860    874       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   875    897       POTENTIAL.
FT   DOMAIN     898    931       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      32     80       EGF-LIKE 1.
FT   DOMAIN      81    132       EGF-LIKE 2.
FT   DOMAIN     133    172       EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     173    221       EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     222    271       EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     272    318       EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     319    367       EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT   SITE       506    508       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID    36     48       BY SIMILARITY.
FT   DISULFID    42     57       BY SIMILARITY.
FT   DISULFID    59     79       BY SIMILARITY.
FT   DISULFID    85     98       BY SIMILARITY.
FT   DISULFID    92    107       BY SIMILARITY.
FT   DISULFID   109    131       BY SIMILARITY.
```

```
FT   DISULFID    137    149       BY SIMILARITY.
FT   DISULFID    143    158       BY SIMILARITY.
FT   DISULFID    160    171       BY SIMILARITY.
FT   DISULFID    177    189       BY SIMILARITY.
FT   DISULFID    183    198       BY SIMILARITY.
FT   DISULFID    200    220       BY SIMILARITY.
FT   DISULFID    226    239       BY SIMILARITY.
FT   DISULFID    233    248       BY SIMILARITY.
FT   DISULFID    250    270       BY SIMILARITY.
FT   DISULFID    276    286       BY SIMILARITY.
FT   DISULFID    280    295       BY SIMILARITY.
FT   DISULFID    297    317       BY SIMILARITY.
FT   DISULFID    323    336       BY SIMILARITY.
FT   DISULFID    330    345       BY SIMILARITY.
FT   DISULFID    347    366       BY SIMILARITY.
FT   CARBOHYD    148    148       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    167    167       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    229    229       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    269    269       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    283    283       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    405    405       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    417    417       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    474    474       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    498    498       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    706    706       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   931 AA;  102129 MW;  52963A667E8B76B5 CRC64;
```

```
  Query Match             4.0%;  Score 93;  DB 1;  Length 931;
  Best Local Similarity   20.4%;  Pred. No. 3.6;
  Matches   44;  Conservative   36;  Mismatches   78;  Indels   58;  Gaps    8;

Qy     206 SMWIIRHRCTSHISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQC 265
           |::|| |  |           :|:::|  | || :|  : ::|               |
Db     646 SLYIISHVGT----------VISLVCLALAIATFLLCRAVQNHN-------TYMHLHLC 687

Qy     266 LIMILDHLNLLNGI--------CSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVD- 316
           : : |  :  | ||        |:      |:  :|   |: | :   : :: :| ·|:
Db     688 VCLFLAKILFLTGIDKTDNQTACAIIAGFLHYLFLACFFWMLVEAVMLFLMVRNLKVVNY 747

Qy     317 ---PNYRFLRYNAF--------VWSTAAIMTGSIYIVNQIWENDPSKWNWLPLVGFIRCS 365
               | : |   ||        |  :|::      : |: | |  :        |||
Db     748 FSSRNIKMLHLCAFGYGLPVLVVIISASVQPRGYGMHNRCWLNTET--------GFI--- 796

Qy     366 VKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRK 401
                 | :: ||   : | |  : | |   :|:
Db     797 --------WSFL-GPVCMIITINSVLLAWTLWVLRQ 823
```

```
RESULT   6
CCSA_MARPO
ID   CCSA_MARPO     STANDARD;      PRT;   320 AA.
AC   P12214;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   CYTOCHROME C BIOGENESIS PROTEIN CCSA.
GN   CCSA.
OS   Marchantia polymorpha (Liverwort).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Embryophyta; Marchantiophyta; Marchantiales;
OC   Marchantiaceae; Marchantia.
OX   NCBI_TaxID=3197;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ohyama K.;
RL   Submitted (OCT-1986) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   COMPLETE GENOME.
RA   Ohyama K., Fukuzawa H., Kohchi T., Shirai H., Sano T., Sano S.,
RA   Umesono K., Shiki Y., Takeuchi M., Chang Z., Aota S., Inokuchi H.,
RA   Ozeki H.;
RT   "Chloroplast gene organization deduced from complete sequence of
RT   liverwort Marchantia polymorpha chloroplast DNA.";
RL   Nature 322:572-574(1986).
CC   -!- FUNCTION: REQUIRED DURING CYTOCHROME BIOGENESIS AT THE STEP OF
CC       HEME ATTACHMENT (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE CCMF/CYCK/CCL1/NRFE/CCSA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04465; CAA28133.1; -.
DR   PIR; A05023; A05023.
DR   PIR; S01516; S01516.
DR   InterPro; IPR002541; -.
DR   Pfam; PF01578; CytC_asm; 1.
KW   Cytochrome c-type biogenesis; Chloroplast.
SQ   SEQUENCE   320 AA;  37170 MW;  6003DFE26300668B CRC64;
```

```
  Query Match             3.9%;  Score 91;  DB 1;  Length 320;
  Best Local Similarity   19.5%;  Pred. No. 1.6;
  Matches   63;  Conservative   43;  Mismatches   87;  Indels  130;  Gaps   19;

Qy     148 IYISDFNY-------FLEEVSIQI----FNKCGLIVWFQDGKF--------------WV 181
           ||   | |       | |:|::|         || |   | |              :
Db      24 IYWGKFLYINIKPITILGEISMKIACFFITTFLLIRWSSSGHFPLSNLYESSMFLSWSFT 83

Qy     182 TVDLFMEKQD-----------YCLYRHNFDS-DFPKSM------------WIIRHRCTS 216
           : | :| :              :  | | :   || |            |:: |
Db      84 LIHLILENKSKNTWLGIITAPSAMLTHGFATLSLPKEMQESVFLVPALQSHWLMMHVTMM 143

Qy     217 HISPGSL---EILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHL 273
           :|  :|    :| ||::   ||        | :|                | ::  :
Db     144 MLSYSTLLCGSLLAITILIITLT--------KQKN----------------LPILTSYF 178

Qy     274 NLLNGICSPAGYSSHFFR-MASNLWLSVISYHTWKVLTSLN--------RVDP-NYRFLR 323
           |          ::|  |: :   :   ::|| | || : :|        :|  :|| :
Db     179 NF--------PFNSFIFKNLLQPMENEILSYKTQKVFSFINFRKWQLIKELDNWSYRVIS 230

Qy     324 YNAFVWSTAAIMTGSIYIVNQIWENDPSKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLA 383
             |   | |::|:::  |: |    | ||| |         : |    |:
Db     231 L-GFPLLTIGILSGAVW-ANEAW---GSYWNWDP--------KETWALITWL-------- 269

Qy     384 LSTFNVAMFALTAIYI--RKVKG 404
                 : |||:  | :||
Db     270 ----------IFAIYLHTRMIKG 282
```

```
RESULT   7
CCSA_CYAPA
ID   CCSA_CYAPA     STANDARD;      PRT;   322 AA.
AC   P48257;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   CYTOCHROME C BIOGENESIS PROTEIN CCSA.
GN   CCSA.
OS   Cyanophora paradoxa.
OG   Cyanelle.
OC   Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
OX   NCBI_TaxID=2762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LB555 / PRINGSHEIM;
RA   Stirewalt V.L., Michalowski C.B., Luffelhardt W., Bohnert H.J.,
RA   Bryant D.A.;
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: REQUIRED DURING CYTOCHROME BIOGENESIS AT THE STEP OF
CC       HEME ATTACHMENT (BY SIMILARITY).
```

```
CC   -!- SIMILARITY: BELONGS TO THE CCMF/CYCK/CCL1/NRFE/CCSA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U30821; AAA81298.1; -.
DR   InterPro; IPR002541; -.
DR   Pfam; PF01578; CytC_asm; 1.
KW   Cytochrome c-type biogenesis; Cyanelle.
SQ   SEQUENCE   322 AA;  37453 MW;  7E0CB2E5012A1838 CRC64;


  Query Match             3.9%;  Score 91;  DB 1;  Length 322;
  Best Local Similarity   18.5%;  Pred. No. 1.6;
  Matches   61;  Conservative   50;  Mismatches   81;  Indels  138;  Gaps   17;

Qy     149 YISDFNYFLEEVS----IQIFNKCGLIV-----------WFQDGKFWV----------- 181
           :|  | :| : :     :: ||  |::           |   | | :
Db      19 FILTFLFFAQAIFKNHWLKTFNFSGMLIANVSIFLLLITRWITAGYFPISNLYESLFFLV 78

Qy     182 ----TVDLFMEKQDYCLYRHN-----FDSDFP---------------------------- 204
               |: ||       ||::|      |:
Db      79 WGLNTIWLF-------LYKNNMRIQLLDNSVSLLLTLLVGFLHFILPQEMREISPLVPAL 131

Qy     205 KSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQ 264
           || |:: |     ||  :|      ||  :|:| :|||                    |:
Db     132 KSNWLMMHVSVMMISYATL------MIGSLLSI-IYLY------------------FLY 165

Qy     265 CLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLSVISYHTWKVLTS-------LNRVDP 317
             :  :|: :| :| : |    :|  : :: |  |:      :::|           |:|
Db     166 RVDKLLNKINAVNLMTSNQNSNSLIYNLSFNTLLNSSKSMEYQILLKDLSLIKLSKRLDN 225

Qy     318 -NYRFLRYNAFVWSTAAIMTGSIYIVNQIWENDPSKWNWLPLVGFIRCSVKDWHPSVWIY 376
            :|| : :  |   |  |: |:::  |: |    | |:| |         : |    |:
Db     226 LSYRLITF-GFPLLTIGIIAGAVW-ANEAW---GSYWSWDP--------KETWALITWL- 271

Qy     377 ISGPSLALSTFNVAMFALTAIYI--RKVKG 404
                       : |||:  | :||
Db     272 ----------------IFAIYLHTRIIKG 284


RESULT   8
STE3_YEAST
ID   STE3_YEAST     STANDARD;      PRT;   470 AA.
AC   P06783;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   PHEROMONE A FACTOR RECEPTOR.
GN   STE3 OR YKL178C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86149315; PubMed=3006051;
RA   Hagen D.C., McCaffrey G., Sprague G.F. Jr.;
RT   "Evidence the yeast STE3 gene encodes a receptor for the peptide
RT   pheromone a factor: gene sequence and implications for the structure
RT   of the presumed receptor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:1418-1422(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Nakayama N., Miyajima A., Arai K.;
RT   "Nucleotide sequences of STE2 and STE3, cell type-specific sterile
RT   genes from Saccharomyces cerevisiae.";
RL   EMBO J. 4:2643-2648(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94205264; PubMed=8154185;
RA   Wiemann S., Voss H., Schwager C., Rupp T., Stegemann J.,
RA   Zimmermann J., Grothues D., Sensen C., Erfle H., Hewitt N.,
RA   Banrevi A., Ansorge W.;
RT   "Sequencing and analysis of 51.6 kilobases on the left arm of
RT   chromosome XI from Saccharomyces cerevisiae reveals 23 open reading
RT   frames including the FAS1 gene.";
RL   Yeast 9:1343-1348(1993).
CC   -!- FUNCTION: RECEPTOR FOR THE PEPTIDE PHEROMONE A FACTOR.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO FAMILY 4 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M12239; AAA35113.1; -.
DR   EMBL; X03011; CAA26795.1; -.
DR   EMBL; X74151; CAA52261.1; -.
DR   EMBL; Z28177; CAA82019.1; -.
DR   PIR; B24670; B24670.
DR   GCRDb; GCR_0197; -.
DR   SGD; S0001661; STE3.
DR   InterPro; IPR001499; -.
DR   Pfam; PF02076; STE3; 1.
DR   PRINTS; PR00899; GPCRSTE3.
KW   Transmembrane; G-protein coupled receptor; Pheromone response.
FT   TRANSMEM      6     23       POTENTIAL.
FT   TRANSMEM     30     53       POTENTIAL.
FT   TRANSMEM     71     98       POTENTIAL.
FT   TRANSMEM    117    134       POTENTIAL.
FT   TRANSMEM    156    183       POTENTIAL.
FT   TRANSMEM    206    228       POTENTIAL.
FT   TRANSMEM    267    285       POTENTIAL.
FT   DOMAIN      300    470       HYDROPHILIC.
SQ   SEQUENCE   470 AA;  53708 MW;  5DBF611076B8CA91 CRC64;


  Query Match             3.9%;  Score 90.5;  DB 1;  Length 470;
  Best Local Similarity   17.5%;  Pred. No. 2.7;
  Matches   61;  Conservative   61;  Mismatches   87;  Indels  139;  Gaps   17;

Qy     118 IYFTYMDLQARVPLTDMA----IIRDEFFDCDEMIYISDFNYFLE--EVSIQIFNKCGLI 171
           ||  :  |:|   | |::    |::|        : :  |:| |:     |  :| |  :
Db      91 IYNLHTILKADSVLPDLSSWTKIVKDLVISLFTPVMVMGFSYLLQVFRYGIARYNGCQNL 150

Qy     172 VWFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIITMI 231
           :       |:|  |:                    :||                 :|| : :
Db     151 L----SPTWITTVLY------------------TMW---------------MLIWSFV 171

Qy     232 CFV-LTIAVYLYIKKLRNVTGKCIVCCI-----VSRF----IQCLIMILDHLNLLNGICS 281
             |  |: :::: || ::|  : |: |       ::||    | | |:|
Db     172 GAVYATLVLFVFYKKRKDV--RDILHCTNSGLNLTRFARLLIFCFIII------------ 217

Qy     282 PAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLR-YNAFVWSTAAIMTGSIY 340
                         | :  :  :  | :|: :| |   ::: :|:|
Db     218 ---------------LVMFPFSVYTFVQDLQQVEGHYTFKNTHSSTIWNT--------- 252

Qy     341 IVNQIWENDPSKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFAL--TAIY 398
               | : || :    |:             ::|:|:    |    | :   ||   : :
Db     253 ----IIKFDPGR----PIY-----------NIWLYVLMSYLVFLIFGLGSDALHMYSKF 292

Qy     399 IRKVKGG--------------------INKFTNEEEGRINCINFDSQ 425
           :| :| |                    :|| :: :| | |  : ||:
Db     293 LRSIKLGFVLDMWKRFIDKNKEKRVGILLNKLSSRKESR-NPFSTDSE 339
```

```
RESULT   9
RRPB_CVMJH
ID   RRPB_CVMJH     STANDARD;      PRT;  2731 AA.
AC   P29982;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (ORF1B).
OS   Murine coronavirus MHV (strain JHM).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Coronaviridae; Coronavirus.
OX   NCBI_TaxID=11144;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91111976; PubMed=1846489;
RA   Lee H.-J., Shieh C.-K., Gorbalenya A.E., Koonin E.V., la Monica N.,
RA   Tuler J., Bagdzhardzhyan A., Lai M.M.C.;
RT   "The complete sequence (22 kilobases) of murine coronavirus gene 1
RT   encoding the putative proteases and RNA polymerase.";
RL   Virology 180:567-582(1991).
CC   -!- FUNCTION: THE RNA DEPENDENT RNA POLYMERASE OF CORONAVIRUSES IS
CC       A MULTIFUNCTIONAL PROTEIN: IT CONTAINS THE ACTIVITIES NECESSARY
CC       FOR THE TRANSCRIPTION OF NEGATIVE STRANDED RNA, LEADER RNA,
CC       SUBGENOMIC MRNAS AND PROGENY VIRION RNA.
CC   -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE
CC       + RNA(N).
CC   -!- MISCELLANEOUS: THIS PROTEIN MAY BE TRANSLATED AS A 1A-1B
CC       POLYPROTEIN BY A RIBOSOMAL FRAMESHIFTING MECHANISM.
CC   -!- SIMILARITY: TO THE AVIAN CORONAVIRUS IBV RNA POLYMERASE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M55148; AAA46458.1; -.
DR   PIR; B36815; VFIHJH.
KW   Transferase; RNA-directed RNA polymerase; Helicase; ATP-binding.
FT   DOMAIN      622    824       POLYMERASE.
FT   DOMAIN      944   1014       CYS/HIS-RICH.
FT   DOMAIN     1218   1506       HELICASE.
FT   NP_BIND    1220   1227       ATP (BY SIMILARITY).
SQ   SEQUENCE   2731 AA;  308834 MW;  73B3C4025244CEE5 CRC64;


  Query Match             3.9%;  Score 90.5;  DB 1;  Length 2731;
  Best Local Similarity   19.2%;  Pred. No. 19;
  Matches   78;  Conservative   37;  Mismatches  119;  Indels  173;  Gaps   21;

Qy      27 CNYYDTVDISYIERQ----NDSYLYDDIEIPASLTGYYEFRQFGDGSITPIEKHLRACVC 82
           || :   |:|  | |    :: |  | :: || :: |||          : ||    :
Db     717 CNGHKIEDLSIRELQKRLYSNVYRADHVD-PAFVSEYYEF----------LNKHFSMIIL 765

Qy      83 SVRPCIRICCPAKNFLANGKCDDGLKEELARFKPYIYFTYMDLQARVPLTDMAIIRDEFF 142
           |                    |||                          :     ||
Db     766 S-------------------DDG--------------------------VVCYNSEFA 778

Qy     143 DCDEMIYISDFN---YFLEEVSIQIFNKCGLIVWFQDGKFWVTVDLFMEKQDYCLYRH-- 197
             :  ||||    |:              |:  : | ||  .|:      ::|  :|
Db     779 SKGYIANISDFQQVLYYQNN-----------VFMSEAKCWVETDIEKGPHEFC-SQHTM 825

Qy     198 --NFDSD-----FPKSMWIIRHRC-TSHISPGSLEILIITMICFVL-------------- 235
               | |     :|    |:  |    :      :||   :   :
Db     826 LVKMDGDEVYLPYPDPSRILGAGCFVDDLLKTDSVLLIERFVSLAIDAYPLVYHENPEYQ 885

Qy     236 -TIAVYL-YIKKLRN---------------------------------------VTGKCI 254
              ||| ||||| |                                          | |:
Db     886 NVFRVYLEYIKKLYNDLGNQILDSISVILSTCDGQKFTDETFYKNMYLRSAVMQSVGACV 945

Qy     255 VC----------CIVSRFIQCL-----IMILDH---LNLLNGICSPAGYSSHFFRMASNL 296
```

```
           ||          ||    : |      :|  ||   |::   :|:  |         : |
Db     946 VCSSQTSLRCGSCIRKPLLCCKCAYDHVMSTDHKYVLSVSPYVCNSPGCD---VNDVTKL 1002

Qy     297 WLSVISYHTWKVLTSLNRVDPNYRF-LRYNAFVWST-AAIMTGSIYI 341
           :|  :||:           | | | |  |  |:       ||| ||
Db    1003 YLGGMSYY-------CEAHKPQYSFKLVMNGMVFGLYKQSCTGSPYI 1042


RESULT  10
TNR1_HUMAN
ID   TNR1_HUMAN     STANDARD;      PRT;   455 AA.
AC   P19438;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (TUMOR NECROSIS FACTOR
DE   BINDING PROTEIN 1) (TBP1) (P60) (TNF-R1) (TNF-RI) (P55) (CD120A).
GN   TNFRSF1A OR TNFR1 OR TNFAR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=90235285; PubMed=2158863;
RA   Schall T.J., Lewis M., Koller K.J., Lee A., Rice G.C., Wong G.H.W.,
RA   Getanaga T., Granger G.A., Lentz R., Raab H., Kohr W.J., Goeddel D.V.;
RT   "Molecular cloning and expression of a receptor for human tumor
RT   necrosis factor.";
RL   Cell 61:361-370(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90235284; PubMed=2158862;
RA   Loetscher H., Pan Y.-C.E., Lahm H.-W., Gentz R., Brockhaus M.,
RA   Tabuchi H., Lesslauer W.;
RT   "Molecular cloning and expression of the human 55 kd tumor necrosis
RT   factor receptor.";
RL   Cell 61:351-359(1990).
RN   [3]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 41-53; 110-124 AND 199-201.
RX   MEDLINE=91006021; PubMed=1698610;
RA   Nophar Y., Kemper O., Brakebusch C., Engelmann H., Zwang R.,
RA   Aderka D., Holtmann H., Wallach D.;
RT   "Soluble forms of tumor necrosis factor receptors (TNF-Rs). The cDNA
RT   for the type I TNF-R, cloned using amino acid sequence data of its
RT   soluble form, encodes both the cell surface and a soluble form of the
RT   receptor.";
RL   EMBO J. 9:3269-3278(1990).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91090841; PubMed=1702293;
RA   Himmler A., Maurer-Fogy I., Kroenke M., Scheurich P., Pfizenmaier K.,
RA   Lantz M., Olsson I., Hauptmann R., Stratowa C., Adolf G.R.;
RT   "Molecular cloning and expression of human and rat tumor necrosis
RT   factor receptor chain (p60) and its soluble derivative, tumor
RT   necrosis factor-binding protein.";
RL   DNA Cell Biol. 9:705-715(1990).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=91017509; PubMed=2170974;
RA   Gray P.W., Barrett K., Chantry D., Turner M., Feldman M.;
RT   "Cloning of human tumor necrosis factor (TNF) receptor cDNA and
RT   expression of recombinant soluble TNF-binding protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:7380-7384(1990).
RN   [6]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92250049; PubMed=1315717;
RA   Fuchs P., Strehl S., Dworzak M., Himmler A., Ambros P.F.;
RT   "Structure of the human TNF receptor 1 (p60) gene (TNFR1) and
RT   localization to chromosome 12p13.";
RL   Genomics 13:219-224(1992).
```

```
RN   [7]
RP   SEQUENCE OF 41-45.
RX   MEDLINE=90110215; PubMed=2153136;
RA   Engelmann H., Novick D., Wallach D.;
RT   "Two tumor necrosis factor-binding proteins purified from human
RT   urine. Evidence for immunological cross-reactivity with cell surface
RT   tumor necrosis factor receptors.";
RL   J. Biol. Chem. 265:1531-1536(1990).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.85 ANGSTROMS) OF 30-211.
RX   MEDLINE=93258809; PubMed=8387891;
RA   Banner D.W., D'Arcy A., Janes W., Gentz R., Schoenfeld H.-J.,
RA   Broger C., Loetscher H., Lesslauer W.;
RT   "Crystal structure of the soluble human 55 kd TNF receptor-human TNF
RT   beta complex: implications for TNF receptor activation.";
RL   Cell 73:431-445(1993).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 41-202.
RX   MEDLINE=97094982; PubMed=8939750;
RA   Naismith J.H., Devine T.Q., Khono H., Sprang S.R.;
RT   "Structures of the extracellular domain of the type I tumor necrosis
RT   factor receptor.";
RL   Structure 4:1251-1262(1996).
CC   -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. THE ADAPTOR MOLECULE FADD
CC       RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING
CC       AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC)
CC       PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION WHICH INITIATES THE
CC       SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE
CC       PROTEASES) MEDIATING APOPTOSIS. CONTRIBUTES TO THE INDUCTION OF
CC       NONCYTOCIDAL TNF EFFECTS INCLUDING ANTI-VIRAL STATE AND ACTIVATION
CC       OF THE ACID SPLINGOMIYELINASE.
CC   -!- SUBUNIT: TNF BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
CC       HOMOTRIMERIZATION. ONCE AGGREGATED THE RECEPTORS DEATH DOMAINS
CC       PROVIDE A NOVEL MOLECULAR INTERFACE THAT INTERACTS SPECIFICALLY
CC       WITH THE DEATH DOMAIN OF TRADD. VARIOUS TRADD-INTERACTING
CC       PROTEINS SUCH AS TRAFS, RIP AND POSSIBLY FADD, ARE RECRUITED TO
CC       TNFR1 COMPLEX BY THEIR ASSOCIATION WITH TRADD. THIS COMPLEX
CC       ACTIVATES AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
CC       NF-KAPPA B SIGNALING.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- DOMAIN: THE DOMAIN THAT INDUCES A-SMASE IS PROBABLY IDENTICAL TO
CC       THE DEATH DOMAIN. THE N-SMASE ACTIVATION DOMAIN (NSD) IS BOTH
CC       NECESSARY AND SUFFICIENT FOR ACTIVATION OF N-SMASE.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC   -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD120a entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd120a.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X55313; CAA39021.1; -.
DR   EMBL; M33294; AAA03210.1; -.
DR   EMBL; M58286; AAA36753.1; -.
DR   EMBL; M63121; AAA36754.1; -.
DR   EMBL; M75866; AAA61201.1; -.
DR   EMBL; M75864; AAA61201.1; JOINED.
DR   EMBL; M75865; AAA61201.1; JOINED.
DR   EMBL; M60275; AAA36756.1; -.
DR   EMBL; A21522; CAA01558.1; -.
DR   PIR; A34899; GQHUT1.
DR   PIR; A35010; A35010.
DR   PIR; S12057; S12057.
DR   PIR; A38208; A38208.
DR   PDB; 1TNR; 31-JUL-94.
DR   PDB; 1NCF; 07-DEC-95.
DR   PDB; 1EXT; 11-JAN-97.
DR   MIM; 191190; -.
DR   InterPro; IPR000488; -.
DR   InterPro; IPR001368; -.
DR   Pfam; PF00020; TNFR_c6; 4.
DR   Pfam; PF00531; death; 1.
DR   PROSITE; PS00652; TNFR_NGFR_1; 3.
DR   PROSITE; PS50050; TNFR_NGFR_2; 3.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Receptor; Transmembrane; Glycoprotein; Repeat; Signal; Apoptosis;
KW   3D-structure.
FT   SIGNAL       1     21
FT   CHAIN       22    455       TUMOR NECROSIS FACTOR RECEPTOR 1.
FT   CHAIN       41    291       TUMOR NECROSIS FACTOR BINDING PROTEIN 1.
FT   DOMAIN      22    211       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   212    234       POTENTIAL.
FT   DOMAIN     235    455       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      43    196       4 X TNFR-CYS.
FT   REPEAT      43     82       TNFR-CYS 1.
FT   REPEAT      83    125       TNFR-CYS 2.
FT   REPEAT     126    166       TNFR-CYS 3.
FT   REPEAT     167    196       TNFR-CYS 4.
FT   DOMAIN     338    348       N-SMASE ACTIVATION DOMAIN (NSD).
FT   DOMAIN     356    441       DEATH DOMAIN.
FT   DISULFID    44     58
FT   DISULFID    59     72
FT   DISULFID    62     81
FT   DISULFID    84     99
FT   DISULFID   102    117
FT   DISULFID   105    125
FT   DISULFID   127    143
FT   DISULFID   146    158
FT   DISULFID   149    166
FT   DISULFID   168    179
FT   DISULFID   182    191
FT   DISULFID   185    195
FT   CARBOHYD    54     54       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   145    145       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   151    151       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT   412    412       MISSING (IN REF. 4).
FT   CONFLICT   443    446       GPAA -> APP (IN REF. 4).
SQ   SEQUENCE   455 AA;  50494 MW;  4CEFBA96D03B8225 CRC64;


  Query Match             3.9%;  Score 90;  DB 1;  Length 455;
  Best Local Similarity   22.3%;  Pred. No. 2.9;
  Matches   56;  Conservative   34;  Mismatches   73;  Indels   88;  Gaps   15;

Qy      37 YIERQNDS----------YLYDDIEIPASLTGYYEFRQFGDGSITPIEKHLRACV-CSVR 85
           ||  ||:|           |||:|   |   |   : |:   || |  | ||| |: ||
Db      49 YIHPQNNSICCTKCHKGTYLYNDCPGPGQDT---DCRECESGSFTASENHLRHCLSCS-- 103

Qy      86 PCIRICCPAKNFLANGKCDDGLKEELARFKPYIYFTYMDLQARVPLTDMAIIRDEFFDCD 145
                         ||    ::|:                 :| ::   : ||    |
Db     104 ---------------KC----RKEM--------------GQVEISSCTVDRDTVCGCR 128

Qy     146 EMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFWVTVDL-FMEKQDYCLYRHNFDSDFP 204
           :    : : ::  |   | || | | :   :|    || |   |||:     |
Db     129 K----NQYRHYWSENLFQCFN-CSLCL---NG----TVHLSCQEKQNTVCTCH------- 169

Qy     205 KSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVYLYIKKLRNV-----TGKCIVCCIV 259
             ::  : | |       : |  | :|          : :: ||     :|  ::  :|
Db     170 AGFFLRENECVS--CSNCKKSLECTKLC----------LPQIENVKGTEDSGTTVLLPLV 217

Qy     260 SRFIQCLIMIL 270
             |  ||: :|
Db     218 IFFGLCLLSLL 228


RESULT 11
YNAI_ECOLI
ID   YNAI_ECOLI     STANDARD;      PRT;   343 AA.
AC   P77253;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
```

```
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   HYPOTHETICAL 38.8 KDA PROTEIN IN MPPA-FNR INTERGENIC REGION.
GN   YNAI.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RA   Aiba H., Baba T., Fujita K., Hayashi K., Honjo A., Horiuchi T.,
RA   Ikemoto K., Inada T., Isono K., Isono S., Itoh T., Kanai K., Kasai H.,
RA   Kashimoto K., Kim S., Kimura S., Kitagawa M., Kitakawa M., Makino K.,
RA   Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y.,
RA   Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA   Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE UPF0003 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000231; AAC74412.1; -.
DR   EMBL; D90771; BAA14923.1; -.
DR   EMBL; D90772; BAA14933.1; -.
DR   EcoGene; EG13924; ynaI.
DR   InterPro; IPR001880; -.
DR   Pfam; PF00924; UPF0003; 1.
DR   PROSITE; PS01246; UPF0003; 1.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM     10     30       POTENTIAL.
FT   TRANSMEM     41     61       POTENTIAL.
FT   TRANSMEM     78     98       POTENTIAL.
FT   TRANSMEM    126    146       POTENTIAL.
FT   TRANSMEM    148    168       POTENTIAL.
SQ   SEQUENCE   343 AA;  38754 MW;  469F0B7909480D13 CRC64;


  Query Match             3.8%;  Score 88.5;  DB 1;  Length 343;
  Best Local Similarity   23.1%;  Pred. No. 2.8;
  Matches   37;  Conservative   34;  Mismatches   68;  Indels   21;  Gaps    6;

Qy     164 IFNKCGLIV----WFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHIS 219
           ||  |  ::    ||: |       ||   | : :|     :      ::| :     |:
Db      16 IFGSCAALILMSFWFRRGNRKRKGFLFHAVQ-FLIYTIIISAVGSIINYVIENYKLKFIT 74

Qy     220 PGSLEILIITMICFVLTIAVYL--------YIKKLRNVTGKCIVCCIVSRFIQCLIMIL- 270
           || :: :  ::|  :||| ::|         ||| |::|   |:  |:   |  ::::|
Db      75 PGVIDFICTSLIAVILTIKLFLLINQFEKQQIKKGRDITSARIMSRIIKITIIVVLVLLY 134

Qy     271 -DHLNL-LNGICSPAGYSSHFFRMA-----SNLWLSVISY 303
            :|  : |:|: :  |       ||     || :  :: |
Db     135 GEHFGMSLSGLLTFGGIGGLAVGMAGKDILSNFFSGIMLY 174


RESULT  12
GLP2_HUMAN
ID   GLP2_HUMAN     STANDARD;      PRT;   553 AA.
AC   O95838;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   GLUCAGON-LIKE PEPTIDE 2 RECEPTOR PRECURSOR (GLP-2 RECEPTOR) (GLP-2-R)
DE   (GLP-2R).
GN   GLP2R.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Stomach;
RX   MEDLINE=99145591; PubMed=9990065;
RA   Munroe D.G., Gupta A.K., Kooshesh F., Vyas T.B., Rizkalla G., Wang H.,
RA   Demchyshyn L., Yang Z.-H., Kamboj R.K., Chen H., McCallum K.,
RA   Sumner-Smith M., Drucker D.J., Crivici A.;
RT   "Prototypic G protein-coupled receptor for the intestinotrophic factor
RT   glucagon-like peptide 2.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:1569-1574(1999).
CC   -!- FUNCTION: THIS IS A RECEPTOR FOR GLUCAGON-LIKE PEPTIDE 2. THE
CC       ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC       ACTIVATE ADENYLYL CYCLASE.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF105367; AAD16895.1; -.
DR   MIM; 603659; -.
DR   InterPro; IPR000832; -.
DR   InterPro; IPR001879; -.
DR   Pfam; PF00002; 7tm_2; 1.
DR   PRINTS; PR00249; GPCRSECRETIN.
DR   PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
DR   PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
DR   PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
DR   PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL       1      ?       POTENTIAL.
FT   CHAIN        ?    553       GLUCAGON-LIKE PEPTIDE 2 RECEPTOR.
FT   DOMAIN       ?    179       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   180    202       1 (POTENTIAL).
FT   DOMAIN     203    210       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   211    230       2 (POTENTIAL).
FT   DOMAIN     231    261       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   262    286       3 (POTENTIAL).
FT   DOMAIN     287    298       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   299    322       4 (POTENTIAL).
FT   DOMAIN     323    337       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   338    363       5 (POTENTIAL).
FT   DOMAIN     364    385       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   386    406       6 (POTENTIAL).
FT   DOMAIN     407    421       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   422    442       7 (POTENTIAL).
FT   DOMAIN     443    550       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    97     97       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   113    113       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   148    148       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   162    162       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   553 AA;  63001 MW;  DA37379DF774A8F4 CRC64;


  Query Match             3.8%;  Score 88;  DB 1;  Length 553;
  Best Local Similarity   19.7%;  Pred. No. 5.2;
```

```
  Matches   57;  Conservative   49;  Mismatches  100;  Indels   84;  Gaps   17;

Qy    155 YFLEEVSIQIFNKCGLIVWFQDGKFWVTV----DLFMEKQDYCLYRHNFDSDFPKSMWII 210
          :: || | : :  |        |  | |:    |:: :  : |   |:|  :      :
Db    124 WWSEESSGRAYRHC-----LAQGT-WQTIENATDIWQDDSE-CSENHSFKQN------VD 170

Qy    211 RHRCTSHISPGSLEILI-----ITMICFVLTIAVYLYIKKL---RNVTGKCIVCCIVSRF 262
          |:   |     :|:::       ::|   | : : |:::||   ||      :    : |
Db    171 RYALLS-----TLQLMYTVGYSFSLISLFLALTLLLFLRKLHCTRNYIHMNLFASFILRT 225

Qy    263 IQCLIM----------------ILDHLNLLNGICSPAGYSSHFFRMASNLWLSV--ISY 303
          :  |:                 : :|: ::  |       |:|  |: ||| |  :
Db    226 LAVLVKDVVFYNSYSKRPDNENGWMSYLSEMSTSCRSVQVLLHYFVGANYLWLLVEGLYL 285

Qy    304 HTWKVLTSL--NRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWENDPSKWNWLPLVGF 361
          ||    | |   |: | |  |      |:        :::|          |      ||
Db    286 HTLLEPTVLPERRLWPRYLLLG-----WAFPV-----LFVV---------PW------GF 320

Qy    362 IR-------CSVKDWHPSVWIYISGPSLALSTFNVAMF-ALTAIYIRKVK 403
          |       |   : :  :|  | || :   | |  :|  :  : | |:|
Db    321 ARAHLENTGCWTTNGNKKIWWIIRGPMMLCVTVNFFIFLKILKLLISKLK 370


RESULT  13
YYO1_CAEEL
ID   YYO1_CAEEL     STANDARD;      PRT;   490 AA.
AC   Q18904;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   PROBABLE G PROTEIN-COUPLED RECEPTOR C56G3.1.
GN   C56G3.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Nelson J.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: NOT KNOWN. PUTATIVE RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U39654; AAA80406.1; -.
DR   WormPep; C56G3.1; CE04283.
DR   InterPro; IPR000276; -.
DR   Pfam; PF00001; 7tm_1; 2.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Hypothetical protein; G-protein coupled receptor; Transmembrane;
KW   Glycoprotein.
FT   DOMAIN        1     55       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     56     76       1 (POTENTIAL).
FT   DOMAIN       77     86       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     87    107       2 (POTENTIAL).
FT   DOMAIN      108    125       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    126    146       3 (POTENTIAL).
FT   DOMAIN      147    166       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    167    187       4 (POTENTIAL).
FT   DOMAIN      188    236       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    237    257       5 (POTENTIAL).
FT   DOMAIN      258    331       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    332    352       6 (POTENTIAL).
FT   DOMAIN      353    374       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    375    395       7 (POTENTIAL).
FT   DOMAIN      396    490       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     17     17       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   490 AA;  56243 MW;  531E1FFAE003CA62 CRC64;


  Query Match             3.7%;  Score 87;  DB 1;  Length 490;
  Best Local Similarity   20.7%;  Pred. No. 5.6;
  Matches   75;  Conservative   45;  Mismatches  105;  Indels  138;  Gaps   19;

Qy     87 CIRICCP-AKNFLA---NGKCDD----GLKEELARFKPY---IYFT-YMDLQARV----P 130
          | |   | | |:|    :| |       |:  | || ||   || | |: :
Db     10 CDRGISPNASNYLTYPFDGLCLQKFFYQLQTSLRRFTPYEEIIYTTVYIIISVAAVIGNG 69

Qy    131 LTDMAIIRDEFFDCDEMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFWVTVDLFMEKQ 190
          |  ||::| :    :  : |   |  |  : : | |     : |
Db     70 LVIMAVVRKKTMRTNRNVLI--LNLALSNLILAITN-----IPF---------------- 106

Qy    191 DYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIITMICFVLTIAV-----YLYIKK 245
            |:  : | :|| | :  :    ::: |||       : |  |||:|     |  :||
Db    107 ---LWLPSIDFEFPYSRFFCKF---ANVLPGS------NIYCSTLTISVMAIDRYYSVKK 154

Qy    246 LRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLSVISYHT 305
          |:          | |   ||   :|                                :|
Db    155 LK----------IASNRKQCFHAVL-----------------------------VSLAI 174

Qy    306 WKVLTSLNRVDPNYRFLRYNAFVWSTAAIM---TGSIYIVNQIWENDPSKWNWLPLVGFI 362
          | |                :|: :|   ::   |  :|:: :|   | |    :   |:
Db    175 WIV-----------------SFILSLPLLLYYETSMLYVMREIRVVDQSGQEVIRSYGWR 217

Qy    363 RC------SVKDWHPSVWIYIS---------GPSLALSTFNVAMFALTAIYIRKVKGGIN 407
          :|       : |   |: : :|          |   || |||  :        | :|   |
Db    218 QCRLVSAGRLPDITQSIQLLMSILQVAFLYIVPLFVLSIFNVKL-------TRFLKTNAN 270

Qy    408 KFT 410
          | :
Db    271 KMS 273


RESULT  14
YDW6_SCHPO
ID   YDW6_SCHPO     STANDARD;      PRT;   535 AA.
AC   O13912;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   HYPOTHETICAL 60.1 KDA PROTEIN C23C11.06C IN CHROMOSOME I.
GN   SPAC23C11.06C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Brown D., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z98559; CAB11159.1; -.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM     55     75       POTENTIAL.
FT   TRANSMEM     82    102       POTENTIAL.
```

```
FT   TRANSMEM     115     135       POTENTIAL.
FT   TRANSMEM     143     163       POTENTIAL.
FT   TRANSMEM     201     221       POTENTIAL.
FT   TRANSMEM     346     366       POTENTIAL.
SQ   SEQUENCE   535 AA;  60124 MW;  A6AE149AAA2929E2 CRC64;


  Query Match             3.7%;  Score 87;  DB 1;  Length 535;
  Best Local Similarity   18.4%;  Pred. No. 6.1;
  Matches   35;  Conservative   31;  Mismatches   66;  Indels   58;  Gaps    7;

Qy    216 SHISPGSLEILIITMICFVLTIAV-YLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLN 274
          ||:||| : :    :   :| : : ::|         : |:| :::  |  |:: |  :
Db     76 SHLSPGFMPVGFTALAAILLAMQIMFVYSPNAPERVTQRIICFLLA--IDVLVVFLSPI- 132

Qy    275 LLNGICSPAGYSSHFFRMASNLWLSVISYHTWKVLTSL------NRVDPNYRFLRYNAFV 328
             :    |: |: |     ||  ::|   | |:| |          |:|  : :  :
Db    133 ----LRHREGWRSNAF----VLWAFLMS--LWIVITDLMLFRQYKEEHPDYDAIAHRGYY 182

Qy    329 WSTAAIMTGSIYIVNQIWENDPSKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFN 388
          ||                                     | || | :    ||  ||
Db    183 WS-----------------------------------WSPSTWPWRQVGSLTASTIL 204

Qy    389 VAMFALTAIY 398
            :  :  |:
Db    205 SVILTILTIH 214


RESULT  15
5H2A_MACMU
ID   5H2A_MACMU     STANDARD;      PRT;   471 AA.
AC   P50128;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   5-HYDROXYTRYPTAMINE 2A RECEPTOR (5-HT-2A) (SEROTONIN RECEPTOR)
DE   (5-HT-2).
GN   HTR2A.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95315240; PubMed=7794950;
RA   Johnson M.P., Baez M., Kursar J.D., Nelson D.L.;
RT   "Species differences in 5-HT2A receptors: cloned pig and rhesus
RT   monkey 5-HT2A receptors reveal conserved transmembrane homology to
RT   the human rather than rat sequence.";
RL   Biochim. Biophys. Acta 1236:201-206(1995).
CC   -!- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR
CC       5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTION
CC       AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THIS RECEPTOR
CC       MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE
CC       A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM. THIS
CC       RECEPTOR IS INVOLVED IN TRACHEAL SMOOTH MUSCLE CONTRACTION,
CC       BRONCHOCONSTRICTION, AND CONTROL OF ALDOSTERONE PRODUCTION.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       STRONGEST TO THE OTHER 5HT-2 SUBTYPE RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S78209; AAB34691.1; -.
DR   GCRDb; GCR_2029; -.
DR   InterPro; IPR000276; -.
DR   InterPro; IPR000455; -.
DR   InterPro; IPR002231; -.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR00516; 5HT2ARECEPTR.
DR   PRINTS; PR01101; 5HTRECEPTOR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family.
FT   DOMAIN         1      75       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      76      99       1 (POTENTIAL).
FT   DOMAIN       100     110       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     111     132       2 (POTENTIAL).
FT   DOMAIN       133     147       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     148     171       3 (POTENTIAL).
FT   DOMAIN       172     191       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     192     215       4 (POTENTIAL).
FT   DOMAIN       216     233       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     234     254       5 (POTENTIAL).
FT   DOMAIN       255     324       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     325     346       6 (POTENTIAL).
FT   DOMAIN       347     362       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     363     384       7 (POTENTIAL).
FT   DOMAIN       385     471       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD       8       8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      38      38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      44      44       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      51      51       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      54      54       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     148     227       BY SIMILARITY.
SQ   SEQUENCE   471 AA;  52603 MW;  300FFC84DE03D462 CRC64;


  Query Match             3.7%;  Score 86.5;  DB 1;  Length 471;
  Best Local Similarity   23.7%;  Pred. No. 5.9;
  Matches   37;  Conservative   24;  Mismatches   60;  Indels   35;  Gaps    6;

Qy    182 TVDLFMEKQDYCLYRHNFDSDFPKSM----WIIRHR------CTSHISPGSLEIL----- 226
          |  |    :|  || ::|:|    :     | :         |   :||  | :|
Db     15 TNSLMQLNEDTRLYSNDFNSGEANTSDAFNWTVESENRTNLSCEGCLSPSCLSLLHLQEK 74

Qy    227 ----IITMICFVLTIAVYLYI-------KKLRNVTGKCIV-CCIVSRFIQCLIMILDHLN 274
              ::| :  :|||| : :         |||:| |   ::   |    :  |:| :  |
Db     75 NWSALLTAVVIILTIAGNILVIMAVSLEKKLQNATNYFLMSLAIADMLLGFLVMPVSMLT 134

Qy    275 LLNG--------ICSPAGYSSHFFRMASNLWLSVIS 302
          :| |        :|:   |    |  || : |  ||
Db    135 ILYGYRWPLPSKLCAVWIYLDVLFSTASIMHLCAIS 170


Search completed: October 30, 2001, 15:29:08
Job time: 504 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 15:19:39 ; Search time 86.82 Seconds
(without alignments)
649.182 Million cell updates/sec

Title: US-09-619-049-855
Perfect score: 2324
Sequence: 1 MRLPWVIFCTVLLLIFTNNS.....NKFTNEEEGRINCINFDSQT 426

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_16:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2324	100.0	426	5	Q9VSE7	Q9vse7 drosophila
2	689.5	29.7	517	5	Q9GT55	Q9gt55 drosophila
3	684.5	29.5	517	5	Q9GT51	Q9gt51 drosophila
4	681.5	29.3	517	5	Q9GT56	Q9gt56 drosophila
5	680.5	29.3	517	5	Q9GT53	Q9gt53 drosophila
6	679.5	29.2	517	5	Q9GT52	Q9gt52 drosophila
7	679.5	29.2	517	5	Q9GT50	Q9gt50 drosophila
8	675.5	29.1	517	5	Q9GT54	Q9gt54 drosophila
9	671	28.9	514	5	O97148	O97148 drosophila
10	658.5	28.3	513	5	Q9W0R4	Q9w0r4 drosophila
11	625.5	26.9	482	5	Q9VRN2	Q9vrn2 drosophila
12	537.5	23.1	362	5	Q9V817	Q9v817 drosophila
13	501.5	21.6	455	5	Q9GT60	Q9gt60 drosophila
14	501.5	21.6	455	5	Q9GT59	Q9gt59 drosophila
15	501.5	21.6	455	5	Q9GN21	Q9gn21 drosophila
16	499.5	21.5	455	5	Q9GT58	Q9gt58 drosophila
17	499.5	21.5	455	5	Q9GNE8	Q9gne8 drosophila
18	498.5	21.5	455	5	Q9GND6	Q9gnd6 drosophila
19	496.5	21.4	455	5	Q9GN78	Q9gn78 drosophila
20	494.5	21.3	455	5	Q9GT57	Q9gt57 drosophila
21	419	18.0	172	5	Q9VS77	Q9vs77 drosophila
22	344.5	14.8	637	5	Q9W0R5	Q9w0r5 drosophila
23	282	12.1	192	5	Q9V818	Q9v818 drosophila
24	195	8.4	513	5	Q9W0R6	Q9w0r6 drosophila
25	158	6.8	492	5	Q9W0V7	Q9w0v7 drosophila
26	120	5.2	3313	11	O88278	O88278 rattus norv
27	118	5.1	1364	4	O75092	O75092 homo sapien
28	118	5.1	3312	4	Q9NYQ7	Q9nyq7 homo sapien
29	116	5.0	1449	5	Q10922	Q10922 caenorhabdi
30	114.5	4.9	676	5	Q9VXD9	Q9vxd9 drosophila
31	111	4.8	401	14	Q9EME5	Q9eme5 amsacta moo
32	110.5	4.8	564	14	O72736	O72736 cowpox viru
33	106.5	4.6	796	5	Q9VAP3	Q9vap3 drosophila
34	105.5	4.5	323	5	Q22927	Q22927 caenorhabdi
35	105	4.5	1209	5	Q21667	Q21667 caenorhabdi
36	104.5	4.5	1014	4	O00406	O00406 homo sapien
37	103.5	4.5	270	5	Q9N364	Q9n364 caenorhabdi
38	102.5	4.4	503	5	Q9N4T9	Q9n4t9 caenorhabdi
39	102.5	4.4	541	4	Q9Y3K0	Q9y3k0 homo sapien
40	102	4.4	3034	11	O35161	O35161 mus musculu
41	102	4.4	3579	5	Q9V5N8	Q9v5n8 drosophila
42	101.5	4.4	326	10	O04714	O04714 arabidopsis
43	100.5	4.3	514	10	Q9LD86	Q9ld86 arabidopsis
44	100	4.3	1218	14	Q9YMN0	Q9ymn0 lymantria d
45	99.5	4.3	307	4	Q9H7Q2	Q9h7q2 homo sapien

ALIGNMENTS

```
RESULT   1
Q9VSE7
ID   Q9VSE7      PRELIMINARY;      PRT;   426 AA.
AC   Q9VSE7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CG7476 PROTEIN.
GN   CG7476 OR CG7674.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
```

```
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003556; AAF50475.1; -.
DR   FlyBase; FBgn0035847; CG7476.
DR   InterPro; IPR000832; -.
SQ   SEQUENCE   426 AA;  49630 MW;  BF11D7EBE82F0561 CRC64;


  Query Match             100.0%;  Score 2324;  DB 5;  Length 426;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-201;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MRLPWVIFCTVLLLIFTNNSNADIPGCNYYDTVDISYIERQNDSYLYDDIEIPASLTGYY 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MRLPWVIFCTVLLLIFTNNSNADIPGCNYYDTVDISYIERQNDSYLYDDIEIPASLTGYY 60

Qy      61 EFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLKEELARFKPYIYF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 EFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLKEELARFKPYIYF 120

Qy     121 TYMDLQARVPLTDMAIIRDEFFDCDEMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TYMDLQARVPLTDMAIIRDEFFDCDEMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFW 180

Qy     181 VTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVY 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 VTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVY 240

Qy     241 LYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLSV 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 LYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLSV 300

Qy     301 ISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWENDPSKWNWLPLVG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWENDPSKWNWLPLVG 360

Qy     361 FIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGINKFTNEEEGRINCI 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 FIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGINKFTNEEEGRINCI 420

Qy     421 NFDSQT 426
           ||||||
Db     421 NFDSQT 426


RESULT   2
Q9GT55
ID   Q9GT55      PRELIMINARY;      PRT;   517 AA.
AC   Q9GT55;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YAK_2;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280584; AAG21992.1; -.
SQ   SEQUENCE   517 AA;  60094 MW;  D7D724AB77322728 CRC64;


  Query Match             29.7%;  Score 689.5;  DB 5;  Length 517;
  Best Local Similarity   35.2%;  Pred. No. 6e-54;
  Matches  160;  Conservative   77;  Mismatches  159;  Indels   59;  Gaps   15;

Qy       1 MRLPWV-----IFCTVLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPA 54
             |:| ||     |   |: |:    :|| || |:||||||||  ::  |:|||:: : :||
Db       1 MKLFWVKRLLRISTVVVTLLLLQRTNAAIPDCDYYDTVDISAAQKLPNESYLFEGLLVPA 60

Qy      55 SLTGYYEFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKC-DDGLKEELAR 113
           :||  |||    | |   ::||:| ||| ::||:| |||  : :    |  |  :|||
Db      61 NLTAEYEFTILPDDSKQKVDKHIRGCVCKLKPCVRFCCPHNHIMDMNVCAGDMTEEELEV 120

Qy     114 FKPYIYFTYMDLQARVPLTDMAIIRDEF----------FDCDEMIYISDFNYFLEEVSI 162
             |::          | | | :::|  |             || |             |:
Db     121 LDPFL---------NVTLDDGSVVRRHFKKELIVQWDLPMSCDGMF-----------SL 159

Qy     163 QIFNKCGLIVWFQDGKFWVTVD-LFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISP- 220
               |      |::| |:   | : : |::|||    |      |: |  | |   |||
Db     160 DNREKTDQYTLFENGSFFRYHDRVTLNKREYCLQHLTFVDGNESSIRIAPHNCL--ISPS 217

Qy     221 --GSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNG 278
             |   ::| :::| |||| |||::|||:|: ||| :| :|  |:  |:::|:   :
Db     218 RMGQTVVMITSLVCMVLTITVYLFVKKLQNLHGKCFMCYMVCLFMAYLLLLLNLWQMSQN 277

Qy     279 ICSPAGYSSHFFRMASNLWLSVISYHTWKVLT----SLNRVDPNYRFLRYNAFVWSTAAI 334
           | ||:  :|| ||: ||||||| | |   :   : ||    :||| || :||  | :
Db     278 FCITAGFLGYFFVMAAFLWLSVISLHLWNTFSGSAHNANRFLSEHRFLAYNTYVWGMAVV 337

Qy     335 MTGSIYIVNQIWENDPSKWNWLPLVGF-IRCSV--KDWHPSVWIYISGPSLALSTFNVAM 391
           :||  |: ::: ||:    :| | :||   | :  : |  |  :|  || : |  ||: |
Db     338 LTGITYLADKVVENE----DWNPRMGFGGHCWICTQSW--SAMLYFYGPMVFLIAFNITM 391

Qy     392 FALTAIYIRKVKGGINKFTNEEEGRINCINFDSQT 426
           | |||  |  ||  | || : :| |   :| | ||
Db     392 FILTANRIIGVKKDIQKFAHRQE-RKQKLNSDKQT 425


RESULT   3
Q9GT51
ID   Q9GT51      PRELIMINARY;      PRT;   517 AA.
AC   Q9GT51;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YAK_5;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280588; AAG21996.1; -.
SQ   SEQUENCE   517 AA;  60032 MW;  7D9726B5D4445981 CRC64;


  Query Match             29.5%;  Score 684.5;  DB 5;  Length 517;
  Best Local Similarity   34.9%;  Pred. No. 1.7e-53;
```

```
  Matches  159;  Conservative   76;  Mismatches  161;  Indels   59;  Gaps   15;

Qy       1 MRLPWV-----IFCTVLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPA 54
           |:| ||      |   |: |:    :|| || |:||||||||  ::  | |||:: : :||
Db       1 MKLFWVKRLLWISTVVVTLLLLQRTNAAIPDCDYYDTVDISAAQKLPNGSYLFEGLLVPA 60

Qy      55 SLTGYYEFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKC-DDGLKEELAR 113
           :|| |||    | |   ::||:| ||| ::||:| |||  : :    |  |  :|||
Db      61 NLTAEYEFTILPDDSKQKVDKHIRGCVCKLKPCVRFCCPHNHIMDMNVCAGDMTEEELEV 120

Qy     114 FKPYIYFTYMDLQARVPLTDMAIIRDEF-----------FDCDEMIYISDFNYFLEEVSI 162
             |::          | | | :::|  |              || |             |:
Db     121 LDPFL---------NVTLDDGSVVRRHFKKELIVQWDLPMSCDGMF-----------SL 159

Qy     163 QIFNKCGLIVWFQDGKFWVTVD-LFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISP- 220
             |      |::| |:    | : : |::|||    |      |: |  | |   |||
Db     160 DNREKTDQYTLFENGSFFRHHDRVTLNKREYCLQHLTFVDGNESSIRIAPHNCL--ISPS 217

Qy     221 --GSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNG 278
             |   ::| :::| |||| |||::|||:|: ||| :| :|  |:  |:::|:   :
Db     218 RMGQTVVMITSLVCMVLTITVYLFVKKLQNLHGKCFMCYMVCLFMAYLLLLLNLWQMSQN 277

Qy     279 ICSPAGYSSHFFRMASNLWLSVISYHTWKVLT----SLNRVDPNYRFLRYNAFVWSTAAI 334
             |  ||:  :|| ||: |||||||  | |   :    : ||    :||| || : |  | :
Db     278 FCITAGFLGYFFVMAAFLWLSVISLHLWNTFSGSAHNANRFLSEHRFLAYNTYAWGMAVV 337

Qy     335 MTGSIYIVNQIWENDPSKWNWLPLVGF-IRCSV--KDWHPSVWIYISGPSLALSTFNVAM 391
           :||  |: ::: ||:    :| | :||   | :  : |  |  :|  || : |  ||: |
Db     338 LTGITYLADKVVENE----DWNPRMGFGGHCWICTQSW--SAMLYFYGPMVFLIAFNITM 391

Qy     392 FALTAIYIRKVKGGINKFTNEEEGRINCINFDSQT 426
           | |||  |  ||  | || : :| |   :| | ||
Db     392 FILTANRIIGVKKDIQKFAHRQE-RKQKLNSDKQT 425


RESULT   4
Q9GT56
ID   Q9GT56      PRELIMINARY;      PRT;   517 AA.
AC   Q9GT56;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BG1013;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280583; AAG21991.1; -.
SQ   SEQUENCE   517 AA;  59954 MW;  839CBF03F93283CF CRC64;


  Query Match             29.3%;  Score 681.5;  DB 5;  Length 517;
  Best Local Similarity   34.9%;  Pred. No. 3.1e-53;
  Matches  159;  Conservative   76;  Mismatches  161;  Indels   59;  Gaps   15;

Qy       1 MRLPWV-----IFCTVLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPA 54
           |:| ||      |   |: |:    :|| || |:||||||||  ::  | |||:: : :||
Db       1 MKLFWVKRLLRISTVVVTLLLLQRTNAAIPDCDYYDTVDISAAQKLPNGSYLFEGLLVPA 60

Qy      55 SLTGYYEFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKC-DDGLKEELAR 113
           :|| |||    | |   ::||:| ||| ::||:| |||  : :    |  |  :|||
Db      61 NLTAEYEFTILPDDSKQKVDKHIRGCVCKLKPCVRFCCPHNHIMDMNVCAGDMTEEELEV 120

Qy     114 FKPYIYFTYMDLQARVPLTDMAIIRDEF-----------FDCDEMIYISDFNYFLEEVSI 162
```

```
             |::          | | | :::|  |              || |             |:
Db     121 LDPFL---------NVTLDDGSVVRRHFKKELIVQWDLPMSCDGMF-----------SL 159

Qy     163 QIFNKCGLIVWFQDGKFWVTVD-LFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISP- 220
             |      |::| |:    | : : |::|||    |      |: |  | |   |||
Db     160 DNREKTDQYTLFENGSFFRHHDRVTLNKREYCLQHLTFVDGNESSIRIAPHNCL--ISPS 217

Qy     221 --GSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNG 278
             |   ::| :::| |||| |||::|||:|: ||| :| :|  |:  |:::|:   :
Db     218 RMGQTVVMISSLVCMVLTITVYLFVKKLQNLHGKCFMCYMVCLFMAYLLLLLNLWQMSQN 277

Qy     279 ICSPAGYSSHFFRMASNLWLSVISYHTWKVLT----SLNRVDPNYRFLRYNAFVWSTAAI 334
             |  ||:  :|| ||: |||||||  | |   :    : ||    :||| || : |  | :
Db     278 FCITAGFLGYFFVMAAFLWLSVISLHLWNTFSGSAHNANRFLSEHRFLAYNTYAWGMAVV 337

Qy     335 MTGSIYIVNQIWENDPSKWNWLPLVGF-IRCSV--KDWHPSVWIYISGPSLALSTFNVAM 391
           :||  |: ::: ||:    :| | :||   | :  : |  |  :|  || : |  ||: |
Db     338 LTGITYLADKVVENE----DWNPRMGFGGHCWICTQSW--SAMLYFYGPMVFLIAFNITM 391

Qy     392 FALTAIYIRKVKGGINKFTNEEEGRINCINFDSQT 426
           | |||  |  ||  | || : :| |   :| | ||
Db     392 FILTANRIIGVKKDIQKFAHRQE-RKQKLNSDKQT 425


RESULT   5
Q9GT53
ID   Q9GT53      PRELIMINARY;      PRT;   517 AA.
AC   Q9GT53;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YAK_30;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280586; AAG21994.1; -.
SQ   SEQUENCE   517 AA;  59945 MW;  5880753D7E9EE545 CRC64;


  Query Match             29.3%;  Score 680.5;  DB 5;  Length 517;
  Best Local Similarity   34.9%;  Pred. No. 3.9e-53;
  Matches  159;  Conservative   75;  Mismatches  162;  Indels   59;  Gaps   15;

Qy       1 MRLPWV-----IFCTVLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPA 54
           |:| ||      |   |: |:    :|| || |:||||||||  ::  | |||:: : :||
Db       1 MKLFWVKRLLRISTXVVTLLLLQRTNAAIPDCDYYDTVDISAAQKLPNGSYLFEGLLVPA 60

Qy      55 SLTGYYEFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKC-DDGLKEELAR 113
           :|| |||    | |   ::||:| ||| ::||:| |||  : :    |  |  :|||
Db      61 NLTAEYEFTILPDDSXQKVDKHIRGCVCXLKPCVRFCCPHNHIMDMNVCAGDMTEEELEV 120

Qy     114 FKPYIYFTYMDLQARVPLTDMAIIRDEF-----------FDCDEMIYISDFNYFLEEVSI 162
             |::          | | | :::|  |              || |             |:
Db     121 LDPFL---------NVTLDDGSVVRRHFKKELIVQWDLPMSCDGMF-----------SL 159

Qy     163 QIFNKCGLIVWFQDGKFWVTVD-LFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISP- 220
             |      |::| |     | : : |::|||    |      |: |  | |   |||
Db     160 DNREKTDQYTLFENGSFXRHHDRVTLNKREYCLQHLTFVDGNESSIRIAPHNCL--ISPS 217

Qy     221 --GSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNG 278
             |   ::| :::| |||| |||::|||:|: ||| :| :|  |:  |:::|:   :
Db     218 RMGQTVVMITSLVCMVLTITVYLFVKKLQNLHGKCFMCYMVCLFMAYLLLLLNLWQMSQN 277
```

```
Qy    279 ICSPAGYSSHFFRMASNLWLSVISYHTWKVLT----SLNRVDPNYRFLRYNAFVWSTAAI 334
          |  ||:  :|| ||: |||||||  |  |    :    : ||    :||| || : |  | :
Db    278 FCITAGFLGYFFVMAAFLWLSVISLHLWNTFSGSAHNANRFLSEHRFLAYNTYXWGMAVV 337

Qy    335 MTGSIYIVNQIWENDPSKWNWLPLVGF-IRCSV--KDWHPSVWIYISGPSLALSTFNVAM 391
          :||  |: ::: ||:    :| | :||   | :  : |  |  :|  || : |  ||: |
Db    338 LTGITYLADKVVENE----DWNPRMGFGGHCWICTQSW--SAMLYFYGPMVFLIAFNITM 391

Qy    392 FALTAIYIRKVKGGINKFTNEEEGRINCINFDSQT 426
          | |||  |  ||  | || : :| |   :| | ||
Db    392 FILTANRIIGVKKDIQKFAHRQE-RKQKLNSDKQT 425


RESULT   6
Q9GT52
ID   Q9GT52      PRELIMINARY;      PRT;   517 AA.
AC   Q9GT52;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YAK_4;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280587; AAG21995.1; -.
SQ   SEQUENCE   517 AA;  59938 MW;  872FC1ACFDE837B6 CRC64;


  Query Match             29.2%;  Score 679.5;  DB 5;  Length 517;
  Best Local Similarity   34.9%;  Pred. No. 4.8e-53;
  Matches  159;  Conservative   75;  Mismatches  162;  Indels   59;  Gaps   15;

Qy      1 MRLPWV-----IFCTVLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPA 54
            |:| ||     |   |: |:    :|| || |:|||||||  ::  | |||:: : :||
Db      1 MKLFWVKRLLRISTVVVTLLLLQRTNAAIPDCDYYDTVDISAAQKLPNGSYLFEGLLVPA 60

Qy     55 SLTGYYEFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKC-DDGLKEELAR 113
          :||  |||    | |   ::||:| ||| ::||:| |||  : :    |  |  :|||
Db     61 NLTAEYEFTILPDDSKQKVDKHIRGCVCKLKPCVRFCCPHNHIMDMNVCAGDMTEEELEV 120

Qy    114 FKPYIYFTYMDLQARVPLTDMAIIRDEF----------FDCDEMIYISDFNYFLEEVSI 162
            |::          | | | :::|  |             || |              |:
Db    121 LDPFL---------NVTLDDGSVVRRHFKKELIVQWDLPMSCDGMF-----------SL 159

Qy    163 QIFNKCGLIVWFQDGKFWVTVD-LFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISP- 220
              |      |::| |    | : : |::|||    |      |: |  | |   |||
Db    160 DNREKTDQYTLFENGSFXRHHDRVTLNKREYCLQHLTFVDGBESSIRIAPHNCL--ISPS 217

Qy    221 --GSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNG 278
            |   ::| :::| |||| |||::|||:|: ||| :| :|  |:  |:::|:   :
Db    218 RMGQTVVMITSLVCMVLTITVYLFVKKLQNLHGKCFMCYMVCLFMAYLLLLLNLWQMSQN 277

Qy    279 ICSPAGYSSHFFRMASNLWLSVISYHTWKVLT----SLNRVDPNYRFLRYNAFVWSTAAI 334
          |  ||:  :|| ||: |||||||  |  |    :    : ||    :||| || : |  | :
Db    278 FCITAGFLGYFFVMAAFLWLSVISLHLWNTFSGSAHNANRFLSEHRFLAYNTYAWGMAVV 337

Qy    335 MTGSIYIVNQIWENDPSKWNWLPLVGF-IRCSV--KDWHPSVWIYISGPSLALSTFNVAM 391
          :||  |: ::: ||:    :| | :||   | :  : |  |  :|  || : |  ||: |
Db    338 LTGITYLADKVVENE----DWNPRMGFGGHCWICTQSW--SAMLYFYGPMVFLIAFNITM 391

Qy    392 FALTAIYIRKVKGGINKFTNEEEGRINCINFDSQT 426
          | |||  |  ||  | || : :| |   :| | ||
Db    392 FILTANRIIGVKKDIQKFAHRQE-RKQKLNSDKQT 425


RESULT   7
Q9GT50
ID   Q9GT50      PRELIMINARY;      PRT;   517 AA.
AC   Q9GT50;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YAK_8;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280589; AAG21997.1; -.
SQ   SEQUENCE   517 AA;  59949 MW;  65C0E549C86ADF06 CRC64;


  Query Match             29.2%;  Score 679.5;  DB 5;  Length 517;
  Best Local Similarity   34.7%;  Pred. No. 4.8e-53;
  Matches  158;  Conservative   77;  Mismatches  161;  Indels   59;  Gaps   15;

Qy      1 MRLPWV-----IFCTVLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPA 54
            |:| ||     |   |: |:    :|| || |:|||||:||  ::  | |||:: : :||
Db      1 MKLFWVKRLLRISTVVVTLLLLQRTNAAIPDCDYYDTVNISAAQKLPNGSYLFEGLLVPA 60

Qy     55 SLTGYYEFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKC-DDGLKEELAR 113
          :||  |||    | |   ::||:| ||| ::||:| |||  : :    |  |  :|||
Db     61 NLTAEYEFTILPDDSXQKVDKHIRGCVCKLKPCVRFCCPHNHIMDMNVCAGDMTEEELEV 120

Qy    114 FKPYIYFTYMDLQARVPLTDMAIIRDEF----------FDCDEMIYISDFNYFLEEVSI 162
            |::          | | | :::|  |             || |              |:
Db    121 LDPFL---------NVTLDDGSVVRRHFKKELIVQWDLPMSCDGMF-----------SL 159

Qy    163 QIFNKCGLIVWFQDGKFWVTVD-LFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISP- 220
              |      |::| |    | : : |::|||    |      |: |  | |   |||
Db    160 DNREKTDQYTLFENGSFFRHHDRVTLNKREYCLQHLTFVDGNESSIRIAPHNCL--ISPS 217

Qy    221 --GSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNG 278
            |   ::| :::| |||| |||::|||:|: ||| :| :|  |:  |:::|:   :
Db    218 RMGQTVVMITSLVCMVLTITVYLFVKKLQNLHGKCFMCYMVCLFMAYLLLLLNLWQMSQN 277

Qy    279 ICSPAGYSSHFFRMASNLWLSVISYHTWKVLT----SLNRVDPNYRFLRYNAFVWSTAAI 334
          |  ||:  :|| ||: |||||||  |  |    :    : ||    :||| || : |  | :
Db    278 FCITAGFLGYFFVMAAFLWLSVISLHLWNTFSGSAHNANRFLSEHRFLAYNTYAWGMAVV 337

Qy    335 MTGSIYIVNQIWENDPSKWNWLPLVGF-IRCSV--KDWHPSVWIYISGPSLALSTFNVAM 391
          :||  |: ::: ||:    :| | :||   | :  : |  |  :|  || : |  ||: |
Db    338 LTGITYLADKVVENE----DWNPRMGFGGHCWICTQSW--SAMLYFYGPMVFLIAFNITM 391

Qy    392 FALTAIYIRKVKGGINKFTNEEEGRINCINFDSQT 426
          | |||  |  ||  | || : :| |   :| | ||
Db    392 FILTANRIIGVKKDIQKFAHRQE-RKQKLNSDKQT 425


RESULT   8
Q9GT54
ID   Q9GT54      PRELIMINARY;      PRT;   517 AA.
AC   Q9GT54;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
```

```
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YAK_23;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280585; AAG21993.1; -.
SQ   SEQUENCE   517 AA;  59958 MW;  FDB3E5E228398AF9 CRC64;


  Query Match             29.1%;  Score 675.5;  DB 5;  Length 517;
  Best Local Similarity   34.7%;  Pred. No. 1.1e-52;
  Matches  158;  Conservative   76;  Mismatches  162;  Indels   59;  Gaps   15;

Qy       1 MRLPWV-----IFCTVLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPA 54
           |:| ||      |   |: |:    :|| || |:|||||:||  ::  | |||:  : :||
Db       1 MKLFWVKRLLRISTVVVTLLLLQRTNAAIPDCDYYDTVNISAAQKLPNGSYLFXGLLVPA 60

Qy      55 SLTGYYEFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKC-DDGLKEELAR 113
           :||  |||    | |   ::||:| ||| ::||:| |||  : :   |  |  :|||
Db      61 NLTAEYEFTILPDDSKQKVDKHIRGCVCKLKPCVRFCCPHNHIMDMNVCAGDMTEEELEV 120

Qy     114 FKPYIYFTYMDLQARVPLTDMAIIRDEF-----------FDCDEMIYISDFNYFLEEVSI 162
             |::          | | | :::|  |             || |             |:
Db     121 LDPFL---------NVTLDDGSVVRRHFKKELIVQWDLPMSCDGMF-----------SL 159

Qy     163 QIFNKCGLIVWFQDGKFWVTVD-LFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISP- 220
               |      |::| |:   | : : |::|||    |      |: |  | |   |||
Db     160 DNREKTDQYTLFENGSFFRHHDRVTLNKREYCLQHLTFVDGNESSIRIAPHNCL--ISPS 217

Qy     221 --GSLEILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNG 278
             |   ::| :::| |||| |||::|||:|: ||| :| :|   |:  |:::|:   :
Db     218 RMGQTVVMISSLVCMVLTITVYLFVKKLQNLHGKCFMCYMVCHFMAYLLLLLNLWQMSQN 277

Qy     279 ICSPAGYSSHFFRMASNLWLSVISYHTWKVLT----SLNRVDPNYRFLRYNAFVWSTAAI 334
           |  ||:  :|| ||: ||||||| | |   :    : ||    :||| || : |  | :
Db     278 FCITAGFLGYFFVMAAFLWLSVISLHLWNTFSGSAHNANRFLSEHRFLAYNTYAWGMAVV 337

Qy     335 MTGSIYIVNQIWENDPSKWNWLPLVGF-IRCSV--KDWHPSVWIYISGPSLALSTFNVAM 391
           :||  |: ::: ||:    :| | :||   | :  : |  |  :|  || : |  ||: |
Db     338 LTGITYLADKVVENE----DWNPRMGFGGHCWICTQSW--SAMLYFYGPMVFLIAFNITM 391

Qy     392 FALTAIYIRKVKGGINKFTNEEEGRINCINFDSQT 426
           | |||  |  ||  | || : :| |   :| | ||
Db     392 FILTANRIIGVKKDIQKFAHRQE-RKQKLNSDKQT 425


RESULT   9
O97148
ID   O97148      PRELIMINARY;      PRT;   514 AA.
AC   O97148;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PUTATIVE G-PROTEIN-COUPLED RECEPTOR.
GN   MTH OR CG6936.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON S;
RX   MEDLINE=99011353; PubMed=9794765;
RA   Lin Y.J., Seroude L., Benzer S.;
RT   "Extended life-span and stress resistance in the Drosophila mutant
RT   methuselah.";
RL   Science 282:943-946(1998).
DR   EMBL; AF109308; AAD16981.1; -.
DR   FlyBase; FBgn0023000; mth.
DR   InterPro; IPR000832; -.
DR   Pfam; PF00002; 7tm_2; 1.
KW   Receptor.
SQ   SEQUENCE   514 AA;  59644 MW;  F2A5370CECCD1D5C CRC64;


  Query Match             28.9%;  Score 671;  DB 5;  Length 514;
  Best Local Similarity   36.2%;  Pred. No. 2.8e-52;
  Matches  160;  Conservative   68;  Mismatches  164;  Indels   50;  Gaps   14;

Qy       7 IFCTVLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPASLTGYYEFRQF 65
           |   :|:::    | |||  |:|:|||||  :: || |||:: : :|| ||| |:||
Db       9 ISTVILVVLVIQKSYADILECDYFDTVDISAAQKLQNGSYLFEGLLVPAILTGEYDFRIL 68

Qy      66 GDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLK-EELARFKPYIYFTYMD 124
             | |   : :|:| ||| ::||:| |||  : : || | | :  ||||   |::
Db      69 PDDSKQKVARHIRGCVCKLKPCVRFCCPHDHIMDNGVCYDNMSDEELAELDPFL------ 122

Qy     125 LQARVPLTDMAIIRDEF-----------FDCDEMIYISDFNYFLEEVSIQIFNKCGLIVW 173
             | | | :: |  |              || | |: :     ||      :|  |
Db     123 ---NVTLDDGSVSRRHFKNELIVQWDLPMPCDGMFYLDN----REEQ-----DKYTL--- 167

Qy     174 FQDGKFWVTVD-LFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISP---GSLEILIIT 229
           |::| |:   | : : |::|||    |       |: |  | |   |  |   ::| :
Db     168 FENGTFFRHFDRVTLRKREYCLQHLTFADGNATSIRIAPHNCL--IVPSITGQTVVMISS 225

Qy     230 MICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHF 289
           :|| |||||||::|||:|: ||| :| :|  |:  | ::||   :    | |||:  :|
Db     226 LICMVLTIAVYLFVKKLQNLHGKCFICYMVCLFMGYLFLLLDLWQISISFCKPAGFLGYF 285

Qy     290 FRMASNLWLSVISYHTWKVLT----SLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQI 345
           | ||:  ||||||  | |          ||    :||| || : |  | ::||   : : |
Db     286 FVMAAFFWLSVISLHLWNTFRGSSHKANRFLFEHRFLAYNTYAWGMAVVLTGITVLADNI 345

Qy     346 WENDPSKWNWLPLVGFI-RCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKG 404
             ||     :| | ||    | :     |  :|  || : |  ||: || |||  |  ||
Db     346 VENQ----DWNPRVGHEGHCWIYTQAWSAMLYFYGPMVFLIAFNITMFILTAKRILGVKK 401

Qy     405 GINKFTNEEEGRINCINFDSQT 426
           |  | : :| |   :| | ||
Db     402 DIQNFAHRQE-RKQKLNSDKQT 422


RESULT  10
Q9W0R4
ID   Q9W0R4      PRELIMINARY;      PRT;   513 AA.
AC   Q9W0R4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   MTH PROTEIN.
GN   MTH OR CG6936.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
```

```
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003468; AAF47379.1; -.
DR   FlyBase; FBgn0023000; mth.
DR   InterPro; IPR000832; -.
DR   Pfam; PF00002; 7tm_2; 1.
SQ   SEQUENCE   513 AA;  59497 MW;  66C8511E97E5E975 CRC64;

  Query Match           28.3%;  Score 658.5;  DB 5;  Length 513;
  Best Local Similarity  36.2%;  Pred. No. 3.7e-51;
  Matches  160;  Conservative   67;  Mismatches  164;  Indels   51;  Gaps   15;

Qy       7 IFCTVLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPASLTGYYEFRQF 65
           |   :|:::    | |||  |:|:|||||  :: || |||:: : :|| ||| |:||
Db       9 ISTVILVVLVIQKSYADILECDYFDTVDISAAQKLQNGSYLFEGLLVPAILTGEYDFRIL 68

Qy      66 GDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLK-EELARFKPYIYFTYMD 124
             | |   : :|:| ||| ::||:| |||  : : || | | :  |||| . |::
Db      69 PDDSKQKVARHIRGCVCKLKPCVRFCCPHDHIMDNGVCYDNMSDEELAELDPFL------ 122

Qy     125 LQARVPLTDMAIIRDEF-----------FDCDEMIYISDFNYFLEEVSIQIFNKCGLIVW 173
               | | :: |  |              || | |: :     ||      :|  |
Db     123 ---NVTLDDGSVSRRHFKNELIVQWDLPMPCDGMFYLDN----REEQ-----DKYTL--- 167

Qy     174 FQDGKFWVTVD-LFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISP---GSLEILIIT 229
           |::| |:   | : : |::|||    |      |: |  | |   | |   |  ::| :
Db     168 FENGTFFRHFDRVTLRKREYCLQHLTFADGNATSIRIAPHNCL--IVPSITGQTVVMISS 225

Qy     230 MICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHF 289
           :|| |||||||::|||:|: ||| :| :|  |:  | ::||   :    | |||   :|
Db     226 LICMVLTIAVYLFVKKLQNLHGKCFICYMVCLFMGYLFLLLDLWQISISFCKPAGL-GYF 284

Qy     290 FRMASNLWLSVISYHTWKVLT----SLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQI 345
           | ||:  |||||| | |          ||    :||| || : |  | ::||   : : |
Db     285 FVMAAFFWLSVISLHLWNTFRGSSHKANRFLFEHRFLAYNTYAWGMAVVLTGITVLADNI 344

Qy     346 WENDPSKWNWLPLVGFI-RCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKG 404
             ||     :| | ||    | :     |  :|  || : |  ||: || |||  |  ||
Db     345 VENQ----DWNPRVGHEGHCWIYTQAWSAMLYFYGPMVFLIAFNITMFILTAKRILGVKK 400

Qy     405 GINKFTNEEEGRINCINFDSQT 426
           |  | : :| |   :| | ||
Db     401 DIQNFAHRQE-RKQKLNSDKQT 421
```

```
RESULT  11
Q9VRN2
ID   Q9VRN2      PRELIMINARY;      PRT;   482 AA.
AC   Q9VRN2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CG17795 PROTEIN.
GN   CG17795.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003565; AAF50761.1; -.
DR   FlyBase; FBgn0035623; CG17795.
DR   InterPro; IPR000276; -.
DR   InterPro; IPR000832; -.
DR   Pfam; PF00002; 7tm_2; 1.
SQ   SEQUENCE   482 AA;  56250 MW;  E14CE324A7E09163 CRC64;

  Query Match           26.9%;  Score 625.5;  DB 5;  Length 482;
  Best Local Similarity  34.0%;  Pred. No. 3.2e-48;
  Matches  149;  Conservative   73;  Mismatches  163;  Indels   53;  Gaps   16;

Qy      11 VLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPASLTGYYEFRQFGDGS 69
             :  |:   :|:|:|  |::|||||||  :|  | ||||: : ||| ||  |||:   :|
Db       9 IYFLLNLQSSSAEIADCSFYDTVDISEGQRLSNGSYLYEGLLIPAHLTAKYEFKLLANGD 68
```

```
Qy      70 ITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLKEELARFKPYIYFTYMDLQARV 129
           :  |:| ||| :| |:| |||  : :  |:|   :  |              :|    |
Db      69 KEQVPSHVRGCVCKLRTCVRFCCPHDHIMDMGECYANMTTEENEL--------LDPMLNV 120

Qy     130 PLTDMAIIRDEF-------FD----CDEMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGK 178
           | | ::::  :        :|    ||:| |: |    ::| ::           |::|:
Db     121 TLDDGSVVQRHYKKELMVQWDLPKPCDDMFYL-DNRDIMDEYTL----------FENGR 168

Qy     179 FWVTVD-LFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISP---GSLEILIITMICFV 234
           | ::::| :|||    |      |: || | |   | |   |   ::| ::|| | |
Db     169 LLRHYDQVYLDKSEYCLQHRTFGEGNNNSIRIIPHNCL--ILPSRTGQTVVMITSLICLV 226

Qy     235 LTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRMAS 294
           ||||||| :||| |: ||| :| ::  |   | ::||   |    |  ||:  :|| ||:
Db     227 LTIAVYLCVKKLMNLEGKCFICYMMCLFFGYLFLLLDLWELSLDFCKAAGFLGYFFVMAA 286

Qy     295 NLWLSVISYHTWKVLT----SLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWENDP 350
           |||:|| | || ||    |:| :    || |: | |:    :||  |: : :  |:
Db     287 FFWLSIISRHYWKCLTNPCASMN-IRSERAFLLYSCFAWAMPLALTGVTYLADNVVNNE- 344

Qy     351 SKWNWLPLVGFI-RCSV--KDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGIN 407
           | | ||    | :  | | |  | :|  || : |  ||: || ||| :|   |  :
Db     345 ---EWQPRVGDEGHCWIYTKSW--SAMVYFYGPMVLLILFNITMFVLTAKHIIDSKRTLR 399

Qy     408 KFTNEEEGRINCINFDSQ 425
           |    ||||  :| | |
Db     400 KIA-RNEGRIQKLNSDKQ 416
```

```
RESULT  12
Q9V817
ID   Q9V817      PRELIMINARY;      PRT;   362 AA.
AC   Q9V817;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CG6536 PROTEIN.
GN   CG6536.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003803; AAF57864.2; -.
DR   FlyBase; FBgn0034219; CG6536.
DR   InterPro; IPR000832; -.
DR   Pfam; PF00002; 7tm_2; 1.
SQ   SEQUENCE   362 AA;  41498 MW;  D92EF609D3CF4623 CRC64;
```

```
  Query Match             23.1%;  Score 537.5;  DB 5;  Length 362;
  Best Local Similarity   32.2%;  Pred. No. 2e-40;
  Matches  127;  Conservative   70;  Mismatches  139;  Indels   59;  Gaps   14;

Qy       6 VIFCTVLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPASLTGYYEFRQ 64
           ::   || |:    |||:||||:::||||||   |  | ||||: : ||| ||  |:::
Db       3 ILLIAVLFLLMP-KSNAEIPGCDFFDTVDISKAPRFSNGSYLYEGLLIPAHLTAEYDYKL 61

Qy      65 FGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKC-DDGLKEELARFKPYIYFTYM 123
             | |    :  |:| | | :|||||  |||    :   ||  |  ::|| :  |::
Db      62 LADDSKEKVASHVRGCACHLRPCIRFCCPQYQKMQKSKCYGDMSEDELNKHDPFV----- 116

Qy     124 DLQARVPLTDMAIIRDEFFD-----------CDEMIYISDFNYFLEEVSIQIFNKCGLI 171
                | |:| :::|  | :             |  | ::   |: |      :|    |:
Db     117 ----NVTLSDGSVVRRHFKEDLIVQSDLAKPGCPRMYFL---NHELPGNEFTLFENGSLL 169

Qy     172 VWFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRC---TSHISPGSLEILII 228
             : |  |:|   |::||:   :|  |   |: |  | |   : |          ::|
Db     170 ------RHWDKVEL--SKREYCVQHLSFKDD---SIRIAPHFCPLSSEHSRTWKTVAIVI 218

Qy     229 TMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSH 288
           ::|| :|||:||||::|||||: ||| :| : | |:    ::|:     :| |  ||:  :
Db     219 SLICIILTISVYLYVEKLRNLHGKCFICYLASLFLGYFFLVLNVWKYSSGFCVTAGFLGY 278

Qy     289 FFRMASNLWLSVISYHTWKVLTS-----LNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVN 343
           |  ||:  |||||  |  ::  |     |:|: |   |  || : |    |||   |  :
Db     279 FSVMAAFFWLSVIGIHL-RIKFSLASNCLHRLLPENPFRAYNLYAWGIPLIMTAITYTAD 337

Qy     344 QIWENDPSKWNWLPLVGFIRCSVKDWHPSVWIYIS 378
           |: :|:  :    | ||  :        : ||| |
Db     338 QVVKNEKLR----PRVGVGK--------NCWIYTS 360
```

```
RESULT  13
Q9GT60
ID   Q9GT60      PRELIMINARY;      PRT;   455 AA.
AC   Q9GT60;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA97_6;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
```

```
DR   EMBL; AF280567; AAG22723.1; -.
SQ   SEQUENCE   455 AA;  52834 MW;  D574F7633619A8A9 CRC64;


  Query Match             21.6%;  Score 501.5;  DB 5;  Length 455;
  Best Local Similarity   30.5%;  Pred. No. 4.5e-37;
  Matches  134;  Conservative   59;  Mismatches  143;  Indels  103;  Gaps   15;

Qy       7 IFCTVLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPASLTGYYEFRQF 65
           |   :|:::    | |||  |:|:|||||  :: || |||:: : :|| ||| |:||
Db       9 ISTVILVVLVIQKSYADILECDYFDTVDISAAQKLQNGSYLFEGLIVPAILTGEYDFRIL 68

Qy      66 GDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLK-EELARFKPYIYFTYMD 124
             | |   : :|:| ||| ::||:| |||  : : || | | :  ||||   |::
Db      69 PDDSKQKVARHIRGCVCKLKPCVRFCCPHDHIMDNGVCYDNMSDEELAELDPFL------ 122

Qy     125 LQARVPLTDMAIIRDEF-----------FDCDEMIYISDFNYFLEEVSIQIFNKCGLIVW 173
           | | | :: |  |                || | |: :     ||      :|  |
Db     123 ---NVTLDDGSVSRRHFKNELIVQWDLPMPCDGMFYLDN----REEQ-----DKYTL--- 167

Qy     174 FQDGKFWVTVD-LFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIITMIC 232
           |::| |:   | : : |::|||    |      |: |  | |            ||:
Db     168 FENGTFFRHFDRVTLRKREYCLQHLTFADGNATSIRIAPHNC----------LIVP--- 213

Qy     233 FVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRM 292
                       ::||: :                                |:  :|| |
Db     214 --------------SITGQTV----------------------------GFLGYFFVM 229

Qy     293 ASNLWLSVISYHTWKVLT----SLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWEN 348
           |:  |||||| | |          ||    :||| || : |  | ::||   : : | ||
Db     230 AAFFWLSVISLHLWNTFRGSSHKANRFLFEHRFLAYNTYAWGMAVVLTGITVLADNIVEN 289

Qy     349 DPSKWNWLPLVGFI-RCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGIN 407
             :| | ||    | :     |  :|  || : |  ||: || |||  |  ||  |
Db     290 Q----DWNPRVGHEGHCWIYTQAWSAMLYFYGPMVFLIAFNITMFILTAKRILGVKKDIQ 345

Qy     408 KFTNEEEGRINCINFDSQT 426
           | : :| |   :| | ||
Db     346 NFAHRQE-RKQKLNSDKQT 363


RESULT  14
Q9GT59
ID   Q9GT59      PRELIMINARY;      PRT;   455 AA.
AC   Q9GT59;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ZIM(H)23;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280574; AAG22730.1; -.
SQ   SEQUENCE   455 AA;  52863 MW;  DC21A1E750D5BBE1 CRC64;


  Query Match             21.6%;  Score 501.5;  DB 5;  Length 455;
  Best Local Similarity   30.5%;  Pred. No. 4.5e-37;
  Matches  134;  Conservative   59;  Mismatches  143;  Indels  103;  Gaps   15;

Qy       7 IFCTVLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPASLTGYYEFRQF 65
           |   :|:::    | |||  |:|:|||||  :: || |||:: : :|| ||| |:||
Db       9 ISTVILVVLVIQKSYADILECDYFDTVDISAAQKLQNGSYLFEGLLVPAILTGEYDFRIL 68

Qy      66 GDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLK-EELARFKPYIYFTYMD 124
             | |   : :|:| ||| ::||:| |||  : : || | | :  ||||   |::
Db      69 PDDSKQKVARHIRGCVCKLKPCVRFCCPHDHIMDNGVCYDNMSDEELAELDPFL------ 122

Qy     125 LQARVPLTDMAIIRDEF-----------FDCDEMIYISDFNYFLEEVSIQIFNKCGLIVW 173
           | | | :: |  |                || | |: :     ||      :|  |
Db     123 ---NVTLDDGSVSRRHFKNELIVQWDLPMPCDGMFYLDN----REEQ-----DKYTL--- 167

Qy     174 FQDGKFWVTVD-LFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIITMIC 232
           |::| |:   | : : |::|||    |      |: |  | |            ||:
Db     168 FENGTFFRHFDRVTLRKREYCLQHLTFADGNATSIRIAPHNC----------LIVP--- 213

Qy     233 FVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRM 292
                       ::||: :                                |:  :|| |
Db     214 --------------SITGQTV----------------------------GFLGYFFVM 229

Qy     293 ASNLWLSVISYHTWKVLT----SLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWEN 348
           |:  |||||| | |          ||    :||| || : |  | ::||   : : | ||
Db     230 AAFFWLSVISLHLWNTFRGSSHKANRFLFEHRFLAYNTYAWGMAVVLTGITVLADNIVEN 289

Qy     349 DPSKWNWLPLVGFI-RCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGIN 407
             :| | ||    | :     |  :|  || : |  ||: || |||  |  ||  |
Db     290 Q----DWNPRVGHEGHCWIYTQAWSAMLYFYGPMVFLIAFNITMFILTAKRILGVKKDIQ 345

Qy     408 KFTNEEEGRINCINFDSQT 426
           | : :| |   :| | ||
Db     346 NFVHRQE-RKQKLNSDKQT 363


RESULT  15
Q9GN21
ID   Q9GN21      PRELIMINARY;      PRT;   455 AA.
AC   Q9GN21;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VARIOUS STRAINS;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280579; AAG22735.1; -.
DR   EMBL; AF280553; AAG22709.1; -.
DR   EMBL; AF280554; AAG22710.1; -.
DR   EMBL; AF280555; AAG22711.1; -.
DR   EMBL; AF280557; AAG22713.1; -.
DR   EMBL; AF280558; AAG22714.1; -.
DR   EMBL; AF280560; AAG22716.1; -.
DR   EMBL; AF280562; AAG22718.1; -.
DR   EMBL; AF280564; AAG22720.1; -.
DR   EMBL; AF280565; AAG22721.1; -.
DR   EMBL; AF280566; AAG22722.1; -.
DR   EMBL; AF280568; AAG22724.1; -.
DR   EMBL; AF280570; AAG22726.1; -.
DR   EMBL; AF280571; AAG22727.1; -.
DR   EMBL; AF280572; AAG22728.1; -.
DR   EMBL; AF280573; AAG22729.1; -.
DR   EMBL; AF280575; AAG22731.1; -.
DR   EMBL; AF280576; AAG22732.1; -.
DR   EMBL; AF280578; AAG22734.1; -.
SQ   SEQUENCE   455 AA;  52834 MW;  E75533A4E1FD196B CRC64;
```

```
  Query Match             21.6%;  Score 501.5;  DB 5;  Length 455;
  Best Local Similarity   30.5%;  Pred. No. 4.5e-37;
  Matches  134;  Conservative   59;  Mismatches  143;  Indels  103;  Gaps   15;

Qy       7 IFCTVLLLIFTNNSNADIPGCNYYDTVDISYIER-QNDSYLYDDIEIPASLTGYYEFRQF 65
           |   :|:::    | |||  |:|:||||||  :: || |||:: : :|| ||| |:||
Db       9 ISTVILVVLVIQKSYADILECDYFDTVDISAAQKLQNGSYLFEGLLVPAILTGEYDFRIL 68

Qy      66 GDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLK-EELARFKPYIYFTYMD 124
            | |   : :|:| ||| ::||:| |||  : : || | | :  ||||   |::
Db      69 PDDSKQKVARHIRGCVCKLKPCVRFCCPHDHIMDNGVCYDNMSDEELAELDPFL------ 122

Qy     125 LQARVPLTDMAIIRDEF-----------FDCDEMIYISDFNYFLEEVSIQIFNKCGLIVW 173
               | | :: |  |              || | |: :     ||       :|  |
Db     123 ---NVTLDDGSVSRRHFKNELIVQWDLPMPCDGMFYLDN----REEQ-----DKYTL--- 167

Qy     174 FQDGKFWVTVD-LFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIITMIC 232
           |::| |:   | : : |::|||    |      |: |  | |           ||:
Db     168 FENGTFFRHFDRVTLRKREYCLQHLTFADGNATSIRIAPHNC----------LIVP--- 213

Qy     233 FVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRM 292
                     ::||: :                            |:  :|| |
Db     214 --------------SITGQTV----------------------------GFLGYFFVM 229

Qy     293 ASNLWLSVISYHTWKVLT----SLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWEN 348
           |:  |||||| | |          ||    :||| || : |  | ::||   : : | ||
Db     230 AAFFWLSVISLHLWNTFRGSSHKANRFLFEHRFLAYNTYAWGMAVVLTGITVLADNIVEN 289

Qy     349 DPSKWNWLPLVGFI-RCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGIN 407
                :| | ||    | :     |  :|  || : |  ||: || |||  |  ||  |
Db     290 Q----DWNPRVGHEGHCWIYTQAWSAMLYFYGPMVFLIAFNITMFILTAKRILGVKKDIQ 345

Qy     408 KFTNEEEGRINCINFDSQT 426
           | : :| |   :| | ||
Db     346 NFAHRQE-RKQKLNSDKQT 363


Search completed: October 30, 2001, 15:27:58
Job time: 499 sec
```

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2001, 11:59:13 ; Search time 2044.91 Seconds
(without alignments)
11618.342 Million cell updates/sec

Title: US-09-619-049-854
Perfect score: 1536
Sequence: 1 atgaggcttccttgggtgat.....ggactctaatgatggattct 1536

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1344157 seqs, 7733874588 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*
44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vi1:*
59: gb_vi2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_ro1:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	673	43.8	46913	68	AC023733	AC023733 Drosophil
	2	673	43.8	83545	65	AC020432	AC020432 Drosophil
c	3	673	43.8	281146	4	AE003556	AE003556 Drosophil
c	4	436	28.4	164659	61	AC010037	AC010037 Drosophil
	5	191	12.4	4616	5	AF053725	AF053725 Drosophil
c	6	21	1.4	400	9	AX034794	AX034794 Sequence
	7	21	1.4	164232	73	AC067792	AC067792 Homo sapi
c	8	20	1.3	5354	15	SCYLL021W	Z73126 S.cerevisia

c	9	20	1.3	6041	14	SCSPA2G X53731 S. cerevisi
c	10	20	1.3	32088	14	SC32KBF X97560 S.cerevisia
c	11	20	1.3	32799	15	SPBC1703 AL136536 S.pombe c
c	12	20	1.3	32814	15	SPAC22A12 Z99295 S.pombe chr
c	13	20	1.3	41025	85	AC003975 AC003975 Human Cos
c	14	20	1.3	103734	79	AL355869 AL355869 Homo sapi
c	15	20	1.3	110000	78	ALI37126_3 Continuation (4 of
	16	20	1.3	120249	82	AP000858 AP000858 Homo sapi
	17	20	1.3	122279	92	HS215D11 AL034417 Human DNA
c	18	20	1.3	156660	63	AC012630 AC012630 Homo sapi
c	19	20	1.3	162975	69	AC025225 AC025225 Homo sapi
	20	20	1.3	164689	80	AL359928 AL359928 Homo sapi
c	21	20	1.3	166905	69	AC025729 AC025729 Homo sapi
c	22	20	1.3	169277	75	AC078940 AC078940 Homo sapi
c	23	20	1.3	175936	84	SPNEU1908 AL449930 Streptoco
	24	20	1.3	178803	71	AC032028 AC032028 Homo sapi
c	25	20	1.3	198566	80	AL356253 AL356253 Homo sapi
	26	20	1.3	209832	83	AP001891 AP001891 Homo sapi
c	27	20	1.3	224567	81	AL512402 AL512402 Homo sapi
c	28	20	1.3	234787	75	AC073780 AC073780 Mus muscu
	29	19	1.2	158	54	DM79A10T Z70968 D. melanoga
c	30	19	1.2	281	7	SSC289757 AJ289757 Sus scrof
c	31	19	1.2	353	97	HUMCOL1AOI M32790 Human alpha
c	32	19	1.2	359	97	HUMCOLA1AM M32798 Human alpha
c	33	19	1.2	525	54	G33109 G33109 Eb25F5T Hum
c	34	19	1.2	977	9	A65495 A65495 Sequence 8
c	35	19	1.2	1076	15	SHY401334 AJ401334 Saccharum
c	36	19	1.2	1329	3	MCABH Z33370 M.capricolu
	37	19	1.2	1817	10	I03653 I03653 Sequence 1
c	38	19	1.2	2252	97	HUMCOL1AZ M55998 Human alpha
c	39	19	1.2	3267	59	HSV1ICP Y00453 Herpes simp
c	40	19	1.2	3347	97	HUMCG1PA1 K01228 Human proal
	41	19	1.2	3755	10	E03025 E03025 DNA encodin
	42	19	1.2	3755	10	E03113 E03113 DNA encodin
	43	19	1.2	3758	59	HS1GB K01760 HSV1 (KOS)
	44	19	1.2	3996	10	E03026 E03026 DNA encodin
	45	19	1.2	3996	10	E03114 E03114 DNA encodin
	46	19	1.2	3997	10	E00357 E00357 DNA encodin
	47	19	1.2	4068	15	SCYGR286C Z73071 S.cerevisia
c	48	19	1.2	4269	6	HROENTI D14038 Halocynthia
c	49	19	1.2	4696	7	AF153062 AF153062 Canis fam
	50	19	1.2	5352	15	SCYGR287C Z73072 S.cerevisia
c	51	19	1.2	5820	13	AF202956 AF202956 Aspergill
	52	19	1.2	6098	59	HS1GBICP M14164 HSV1 (strai
c	53	19	1.2	6728	93	HSPPA1ICO Z74615 H.sapiens m
	54	19	1.2	9731	95	RATMIBP1 D37951 Rattus norv
	55	19	1.2	9756	59	HS1GLYB M21629 Herpes simp
	56	19	1.2	13452	93	HSGOLF03 U55182 Human G pro
c	57	19	1.2	18609	88	AF017178 AF017178 Homo sapi
	58	19	1.2	25560	6	CEZK809 Z68303 Caenorhabdi
c	59	19	1.2	32968	15	SPAC1952 AL109820 S.pombe c
c	60	19	1.2	35400	14	SCCHVIIRA X94332 S.cerevisia
	61	19	1.2	37004	6	CEY53H1C AL117201 Caenorhab
c	62	19	1.2	55469	90	AL445490 AL445490 Human DNA
c	63	19	1.2	55720	78	AC090390 AC090390 Homo sapi
	64	19	1.2	62236	77	AC090297 AC090297 Homo sapi
	65	19	1.2	78242	82	AP000483 AP000483 Homo sapi
c	66	19	1.2	87581	14	ATT7H20 AL162508 Arabidops
c	67	19	1.2	88877	93	HSJ190I16 AL121714 Human DNA
	68	19	1.2	100469	13	ATF19H22 AL035679 Arabidops
	69	19	1.2	103554	83	CEY53H1_3 Continuation (4 of
	70	19	1.2	107244	83	CEY24F12 AL022277 Caenorhab
c	71	19	1.2	108360	59	HS1ULR D10879 Herpes simp
c	72	19	1.2	127824	85	AC004822 AC004822 Homo sapi
c	73	19	1.2	131694	61	AC009619 AC009619 Homo sapi
c	74	19	1.2	136822	71	AC034222 AC034222 Homo sapi
	75	19	1.2	144165	85	AC002127 AC002127 Human BAC
	76	19	1.2	148692	78	AC090507 AC090507 Homo sapi
c	77	19	1.2	149394	82	AP000828 AP000828 Homo sapi
	78	19	1.2	149977	79	AL355313 AL355313 Homo sapi
c	79	19	1.2	150237	80	AL357335 AL357335 Homo sapi
c	80	19	1.2	151886	81	AL391812 AL391812 Homo sapi
c	81	19	1.2	152261	59	HE1CG X14112 Herpes simp
	82	19	1.2	155608	71	AC037437 AC037437 Homo sapi
	83	19	1.2	156492	66	AC020582 AC020582 Homo sapi
c	84	19	1.2	157801	69	AC025380 AC025380 Homo sapi
c	85	19	1.2	158045	86	AC007432 AC007432 Homo sapi
	86	19	1.2	159309	79	AL157896 AL157896 Homo sapi
c	87	19	1.2	162448	72	AC053530 AC053530 Homo sapi
c	88	19	1.2	164727	82	AL590240 AL590240 Homo sapi
c	89	19	1.2	164808	93	HSDJ293L8 AL078594 Human DNA
	90	19	1.2	166034	73	AC067913 AC067913 Homo sapi
c	91	19	1.2	166217	81	AL391706 AL391706 Homo sapi
	92	19	1.2	167764	75	AC073870 AC073870 Homo sapi
c	93	19	1.2	170298	80	AL360217 AL360217 Homo sapi
	94	19	1.2	171021	73	AC068121 AC068121 Homo sapi
c	95	19	1.2	171195	77	AC084806 AC084806 Homo sapi
	96	19	1.2	171398	24	AC022060 Ac022060 Homo sapi
	97	19	1.2	172043	81	AL391813 AL391813 Homo sapi
	98	19	1.2	172064	90	AL445195 AL445195 Human DNA
	99	19	1.2	172209	75	AC074253 AC074253 Homo sapi
	100	19	1.2	173654	82	AP001269 AP001269 Homo sapi
	101	19	1.2	173890	82	AL590097 AL590097 Homo sapi
c	102	19	1.2	173921	88	AC084759 AC084759 Homo sapi
	103	19	1.2	174065	64	AC016731 AC016731 Homo sapi
	104	19	1.2	175483	65	AC018553 AC018553 Homo sapi
	105	19	1.2	176781	87	AC010081 AC010081 Homo sapi
c	106	19	1.2	178004	82	AL590070 AL590070 Homo sapi
c	107	19	1.2	179993	64	AC015879 AC015879 Homo sapi
c	108	19	1.2	180756	78	AL137789 AL137789 Homo sapi
	109	19	1.2	182180	62	AC012365 AC012365 Homo sapi
c	110	19	1.2	182305	73	AC068918 AC068918 Homo sapi
	111	19	1.2	184012	87	AC017080 AC017080 Homo sapi
c	112	19	1.2	184034	68	AC024156 AC024156 Homo sapi
c	113	19	1.2	184251	75	AC073850 AC073850 Homo sapi
	114	19	1.2	184559	73	AC069017 AC069017 Mus muscu
c	115	19	1.2	185492	69	AC025339 AC025339 Homo sapi
	116	19	1.2	185613	76	AC079796 AC079796 Homo sapi
c	117	19	1.2	186607	89	AF131215 AF131215 Homo sapi
c	118	19	1.2	186974	83	AP003071 AP003071 Homo sapi
	119	19	1.2	187540	75	AC073970 AC073970 Homo sapi
c	120	19	1.2	188165	91	CNS00003 AL049781 Human chr
c	121	19	1.2	188608	91	CNS00005 AL049779 Human chr
	122	19	1.2	189397	72	AC051643 AC051643 Homo sapi
	123	19	1.2	193858	64	AC016760 AC016760 Homo sapi
c	124	19	1.2	194781	65	AC019124 AC019124 Homo sapi
	125	19	1.2	198987	13	ATCHRIV90 AL161594 Arabidops
c	126	19	1.2	199951	80	AL356475 AL356475 Homo sapi
c	127	19	1.2	200000	60	AC007118 AC007118 Homo sapi
c	128	19	1.2	200019	68	AC024111 AC024111 Mus muscu
	129	19	1.2	200165	76	AC079904 AC079904 Homo sapi
	130	19	1.2	200484	66	AC020722 AC020722 Homo sapi
	131	19	1.2	200498	82	AP000944 AP000944 Homo sapi
	132	19	1.2	200836	62	AC012360 AC012360 Homo sapi
	133	19	1.2	201015	61	AC010075 AC010075 Homo sapi
	134	19	1.2	212093	64	AC015909 AC015909 Homo sapi
c	135	19	1.2	215811	73	AC069025 AC069025 Homo sapi
	136	19	1.2	224321	74	AC073211 AC073211 Homo sapi
c	137	18	1.2	200	54	G54643 G54643 AffySTS0001
c	138	18	1.2	200	54	G59944 G59944 SHGC-130560
c	139	18	1.2	303	97	HSZ80632 Z80632 H.sapiens B
c	140	18	1.2	321	94	MMCOL1A1 X57981 Mouse gene
	141	18	1.2	551	94	AF093700 AF093700 Mus muscu
	142	18	1.2	820	53	CNS0761J AL430829 clone XBA
c	143	18	1.2	890	1	AB039442 AB039442 Pseudomon
c	144	18	1.2	890	1	AB039444 AB039444 Pseudomon
c	145	18	1.2	890	1	AB039470 AB039470 Pseudomon
c	146	18	1.2	890	1	AB039471 AB039471 Pseudomon
c	147	18	1.2	890	1	AB039472 AB039472 Pseudomon
c	148	18	1.2	967	53	CNS0762T AL430875 clone XBA
c	149	18	1.2	993	53	CNS07347 AL427037 clone BA0
	150	18	1.2	1148	9	AX053058 AX053058 Sequence
	151	18	1.2	1148	95	RRAJ5161 AJ005161 Rattus ra
c	152	18	1.2	1205	94	MMC1AI U03419 Mus musculu
c	153	18	1.2	1455	8	AF121350 AF121350 Gallus ga
c	154	18	1.2	1570	94	MMCOLA X15896 M.musculus

c 155	18	1.2	1570	94	MMCOLA13	X06753 Mouse pro-a
156	18	1.2	1912	85	AB050421	AB050421 Macaca fa
c 157	18	1.2	1988	94	MMU36842	U36842 Mus musculu
c 158	18	1.2	2010	95	S67530S2	S67482 COL1A1-type
c 159	18	1.2	2013	94	BC003198	BC003198 Mus muscu
160	18	1.2	2095	95	RNCLR	X70658 R.norvegicu
c 161	18	1.2	2716	15	YSCSEC9SNA	L34336 Saccharomyc
162	18	1.2	2928	95	S43408	S43408 endopeptida
c 163	18	1.2	3001	89	AK025001	AK025001 Homo sapi
164	18	1.2	3180	15	SCYGR009C	Z72794 S.cerevisia
165	18	1.2	3185	95	RATCRLR	L27487 Rat calcito
166	18	1.2	3209	12	AF048982	AF048982 Arabidops
c 167	18	1.2	4589	94	MMU08020	U08020 Mus musculu
c 168	18	1.2	4636	94	MMU50767	U50767 Mus musculu
c 169	18	1.2	5721	95	RNCA1T1	Z78279 R.norvegicu
170	18	1.2	6286	94	AF246632	AF246632 Mus Muscu
c 171	18	1.2	10668	3	U32844	U32844 Haemophilus
c 172	18	1.2	11914	1	AE004259	AE004259 Vibrio ch
c 173	18	1.2	16422	4	AE001421	AE001421 Plasmodiu
174	18	1.2	24008	64	AC017420	AC017420 Drosophil
c 175	18	1.2	24745	78	AC090847	AC090847 Homo sapi
176	18	1.2	25562	63	AC013984	AC013984 Drosophil
c 177	18	1.2	28687	6	CELF42H10	L08403 C. elegans
178	18	1.2	34107	6	CEF02H6	Z82265 Caenorhabdi
179	18	1.2	35165	6	CELC56C10	U29488 Caenorhabdi
c 180	18	1.2	35828	85	AC005347	AC005347 Homo sapi
c 181	18	1.2	36342	85	AC005364	AC005364 Homo sapi
182	18	1.2	37443	97	U73627	U73627 Human Chrom
183	18	1.2	39329	85	AC004091	AC004091 Human Cos
c 184	18	1.2	40229	5	CBRG42J05	AC084598 Caenorhab
185	18	1.2	40552	4	AC084469	AC084469 Caenorhab
186	18	1.2	41507	63	AC014235	AC014235 Drosophil
187	18	1.2	42696	6	CELC45G7	AF067611 Caenorhab
c 188	18	1.2	43947	6	CELC04C3	AF038606 Caenorhab
c 189	18	1.2	47428	9	AX059539	AX059539 Sequence
c 190	18	1.2	52863	65	AC017834	AC017834 Drosophil
191	18	1.2	55742	77	AC087529	AC087529 Homo sapi
192	18	1.2	58003	67	AC023289	AC023289 Homo sapi
193	18	1.2	62856	64	AC017153	AC017153 Drosophil
194	18	1.2	65177	13	AP001313	AP001313 Arabidops
195	18	1.2	65720	78	AC091037	AC091037 Homo sapi
196	18	1.2	67356	90	AL161658	AL161658 Human DNA
197	18	1.2	71549	86	AC008476	AC008476 Homo sapi
198	18	1.2	72952	90	AL357508	AL357508 Human DNA
c 199	18	1.2	75586	66	AC021444	AC021444 Mus muscu
200	18	1.2	77424	12	AC015449	AC015449 Arabidops
c 201	18	1.2	78631	93	HSAC000057	AC000057 Human BAC
202	18	1.2	78983	70	AC027245	AC027245 Homo sapi
c 203	18	1.2	79714	80	AL357292	AL357292 Homo sapi
204	18	1.2	80025	90	AL391355	AL391355 Human DNA
c 205	18	1.2	80196	15	T32N4	AF162444 Arabidops
206	18	1.2	80376	4	AC004442	AC004442 Drosophil
207	18	1.2	80442	15	T24D18	AC010924 Arabidops
c 208	18	1.2	80452	62	AC011433	AC011433 Homo sapi
209	18	1.2	83735	63	AC012619	AC012619 Homo sapi
210	18	1.2	86152	86	AC008585	AC008585 Homo sapi
211	18	1.2	87924	89	AL137796	AL137796 Human DNA
212	18	1.2	89083	78	AF170802	AF170802 Homo sapi
c 213	18	1.2	90090	77	AC087864	AC087864 Homo sapi
c 214	18	1.2	90180	86	AC008176	AC008176 Homo sapi
215	18	1.2	90707	12	AC027032	AC027032 Arabidops
c 216	18	1.2	93368	89	AL133402	AL133402 Human DNA
c 217	18	1.2	94995	87	AC008893	AC008893 Homo sapi
c 218	18	1.2	95310	13	ATAC011437	AC011437 Arabidops
c 219	18	1.2	96293	79	AL354987	AL354987 Homo sapi
220	18	1.2	97805	75	AC074157	AC074157 Mus muscu
221	18	1.2	98491	94	AF132039	AF132039 Mus muscu
c 222	18	1.2	99406	65	AC019509	AC019509 Drosophil
c 223	18	1.2	99814	13	ATF8J2	AL132969 Arabidops
c 224	18	1.2	100457	66	AC020933	AC020933 Homo sapi
c 225	18	1.2	100806	12	AC015450	AC015450 Arabidops
226	18	1.2	101151	78	AL139148	AL139148 Homo sapi
c 227	18	1.2	102757	90	AL353698	AL353698 Human DNA
c 228	18	1.2	104416	65	AC020324	AC020324 Drosophil
229	18	1.2	104939	88	AC035147	AC035147 Homo sapi
230	18	1.2	105191	90	AL161655	AL161655 Human DNA
c 231	18	1.2	105689	12	AC005824	AC005824 Arabidops
232	18	1.2	105998	78	AL133290	AL133290 Homo sapi
c 233	18	1.2	106300	86	AC008602	AC008602 Homo sapi
234	18	1.2	106991	64	AC016960	AC016960 Homo sapi
c 235	18	1.2	107499	89	AL137804	AL137804 Human DNA
236	18	1.2	109677	64	AC016546	AC016546 Homo sapi
c 237	18	1.2	110000	62	AC011329_0	AC011329 Homo sapi
c 238	18	1.2	110000	80	AL360016_0	AL360016 Homo sapi
c 239	18	1.2	111126	89	AF321237	AF321237 Homo sapi
240	18	1.2	111234	87	AC013451	AC013451 Homo sapi
c 241	18	1.2	112392	85	AC004028	AC004028 Human PAC
c 242	18	1.2	115262	88	AC025464	AC025464 Homo sapi
243	18	1.2	117396	61	AC010015	AC010015 Drosophil
c 244	18	1.2	118355	61	AC010431	AC010431 Homo sapi
245	18	1.2	120314	89	AL133338	AL133338 Human DNA
c 246	18	1.2	121084	61	AC010791	AC010791 Homo sapi
247	18	1.2	122952	82	AP000593	AP000593 Homo sapi
248	18	1.2	123423	94	AC006949	AC006949 Mus muscu
c 249	18	1.2	126306	81	AL391242	AL391242 Homo sapi
c 250	18	1.2	126368	80	AL356272	AL356272 Homo sapi
251	18	1.2	129757	13	ATF2K15	AL132956 Arabidops
252	18	1.2	130491	97	HUAC003977	AC003977 Homo sapi
c 253	18	1.2	131061	89	AF121897	AF121897 Homo sapi
c 254	18	1.2	133361	77	AC087263	AC087263 Mus muscu
255	18	1.2	133560	71	AC040986	AC040986 Homo sapi
256	18	1.2	133609	78	AL139098	AL139098 Homo sapi
c 257	18	1.2	133801	85	AC004585	AC004585 Homo sapi
c 258	18	1.2	134059	64	AC016284	AC016284 Homo sapi
c 259	18	1.2	134142	80	AL359976	AL359976 Homo sapi
260	18	1.2	135808	81	AL513013	AL513013 Homo sapi
261	18	1.2	136810	62	AC012172	AC012172 Homo sapi
262	18	1.2	136883	67	AC022926	AC022926 Homo sapi
c 263	18	1.2	137176	85	AC005017	AC005017 Homo sapi
c 264	18	1.2	137404	79	AL162402	AL162402 Homo sapi
c 265	18	1.2	140069	79	AL355005	AL355005 Mus muscu
266	18	1.2	140721	66	AC020591	AC020591 Homo sapi
267	18	1.2	141901	76	AC084028	AC084028 Homo sapi
268	18	1.2	144079	68	AC024342	AC024342 Homo sapi
269	18	1.2	144645	88	AC026273	AC026273 Homo sapi
270	18	1.2	145963	70	AC027390	AC027390 Homo sapi
c 271	18	1.2	146120	94	AC023789	AC023789 Mus muscu
c 272	18	1.2	146248	63	AC013677	AC013677 Homo sapi
c 273	18	1.2	147181	81	AL583860	AL583860 Homo sapi
c 274	18	1.2	148416	68	AC024655	AC024655 Homo sapi
c 275	18	1.2	148588	75	AC073827	AC073827 Mus muscu
276	18	1.2	148641	92	HS1119D9	AL031652 Human DNA
277	18	1.2	148642	92	HS580G13	AL034426 Human DNA
278	18	1.2	150169	77	AC087251	AC087251 Rattus no
c 279	18	1.2	150865	67	AC022666	AC022666 Homo sapi
c 280	18	1.2	151227	62	AC011267	AC011267 Homo sapi
c 281	18	1.2	152626	66	AC021841	AC021841 Homo sapi
282	18	1.2	153106	73	AC068632	AC068632 Homo sapi
283	18	1.2	153125	67	AC021940	AC021940 Homo sapi
c 284	18	1.2	153392	89	AL133548	AL133548 Human DNA
285	18	1.2	153511	61	AC010557	AC010557 Drosophil
286	18	1.2	153873	80	AL365270	AL365270 Homo sapi
c 287	18	1.2	153925	64	AC016424	AC016424 Homo sapi
c 288	18	1.2	154589	86	AC006557	AC006557 Homo sapi
289	18	1.2	155065	93	HSJ659I19	AL122003 Human DNA
290	18	1.2	155291	83	AP002427	AP002427 Homo sapi
c 291	18	1.2	155473	62	AC012252	AC012252 Homo sapi
292	18	1.2	155486	81	AL513364	AL513364 Homo sapi
c 293	18	1.2	156129	90	AL359714	AL359714 Human DNA
294	18	1.2	156660	63	AC012630	AC012630 Homo sapi
295	18	1.2	156741	79	AL355867	AL355867 Homo sapi
296	18	1.2	157073	85	AC005027	AC005027 Homo sapi
297	18	1.2	157213	81	AL449106	AL449106 Homo sapi
c 298	18	1.2	157289	66	AC021447	AC021447 Homo sapi
299	18	1.2	157308	76	AC079351	AC079351 Homo sapi
c 300	18	1.2	157386	61	AC010626	AC010626 Homo sapi

c	301	18	1.2	157434	71	AC034151 AC034151 Homo sapi
	302	18	1.2	158052	88	AC027612 AC027612 Homo sapi
	303	18	1.2	158298	69	AC025038 AC025038 Homo sapi
c	304	18	1.2	159698	81	AL390842 AL390842 Homo sapi
	305	18	1.2	160479	61	AC010539 AC010539 Homo sapi
	306	18	1.2	161198	64	AC015867 AC015867 Homo sapi
	307	18	1.2	161503	60	AC007994 AC007994 Mus muscu
c	308	18	1.2	161694	71	AC037452 AC037452 Homo sapi
c	309	18	1.2	161907	82	AP001810 AP001810 Homo sapi
	310	18	1.2	161973	97	HUAC002038 AC002038 Homo sapi
c	311	18	1.2	162496	71	AC044846 AC044846 Mus muscu
	312	18	1.2	162782	80	AL359955 AL359955 Homo sapi
c	313	18	1.2	162902	70	AC026459 AC026459 Homo sapi
c	314	18	1.2	162978	88	AC021070 AC021070 Homo sapi
c	315	18	1.2	163332	85	AC004821 AC004821 Homo sapi
c	316	18	1.2	164232	73	AC067792 AC067792 Homo sapi
	317	18	1.2	164335	76	AC079932 AC079932 Homo sapi
c	318	18	1.2	164469	73	AC068518 AC068518 Homo sapi
	319	18	1.2	164485	69	AC026106 AC026106 Homo sapi
c	320	18	1.2	164794	66	AC021331 AC021331 Homo sapi
c	321	18	1.2	164868	87	AC016045 AC016045 Homo sapi
c	322	18	1.2	164927	65	AC018629 AC018629 Homo sapi
	323	18	1.2	165213	80	AL356576 AL356576 Homo sapi
	324	18	1.2	166105	69	AC024972 AC024972 Homo sapi
	325	18	1.2	166438	66	AC021838 AC021838 Homo sapi
	326	18	1.2	167765	78	AF182108 AF182108 Homo sapi
c	327	18	1.2	168248	62	AC012002 AC012002 Homo sapi
c	328	18	1.2	168270	78	AL138926 AL138926 Homo sapi
c	329	18	1.2	168522	68	AC024302 AC024302 Homo sapi
	330	18	1.2	169090	60	AC008432 AC008432 Homo sapi
	331	18	1.2	169350	68	AC024332 AC024332 Homo sapi
	332	18	1.2	169693	92	HS57A13 Z83848 Human DNA s
	333	18	1.2	169813	72	AC055735 AC055735 Homo sapi
c	334	18	1.2	170172	83	AP002892 AP002892 Homo sapi
c	335	18	1.2	170328	68	AC024169 AC024169 Homo sapi
c	336	18	1.2	170348	88	AC068312 AC068312 Homo sapi
	337	18	1.2	170545	61	AC010098 AC010098 Homo sapi
	338	18	1.2	170970	66	AC021564 AC021564 Homo sapi
	339	18	1.2	170996	63	AC012686 AC012686 Homo sapi
c	340	18	1.2	171025	80	AL356302 AL356302 Homo sapi
c	341	18	1.2	171398	24	AC022060 Ac022060 Homo sapi
	342	18	1.2	171548	68	AC024732 AC024732 Homo sapi
	343	18	1.2	171660	64	AC017103 AC017103 Homo sapi
c	344	18	1.2	172056	65	AC019345 AC019345 Homo sapi
c	345	18	1.2	172286	61	AC010387 AC010387 Homo sapi
	346	18	1.2	172328	70	AC027233 AC027233 Homo sapi
c	347	18	1.2	172332	61	AC009122 AC009122 Homo sapi
c	348	18	1.2	172657	90	AL139009 AL139009 Human DNA
	349	18	1.2	173083	63	AC013277 AC013277 Homo sapi
	350	18	1.2	173677	78	AC091290 AC091290 Mus muscu
	351	18	1.2	173705	88	AC073175 AC073175 Homo sapi
c	352	18	1.2	173872	61	AC009584 AC009584 Homo sapi
	353	18	1.2	173921	88	AC084759 AC084759 Homo sapi
c	354	18	1.2	174054	63	AC015160 AC015160 Drosophil
	355	18	1.2	174150	82	AP001004 AP001004 Homo sapi
	356	18	1.2	174500	4	AC009202 AC009202 Drosophil
c	357	18	1.2	174769	66	AC021674 AC021674 Homo sapi
c	358	18	1.2	174899	77	AC084277 AC084277 Homo sapi
	359	18	1.2	175582	64	AC016977 AC016977 Homo sapi
	360	18	1.2	175772	77	AC084381 AC084381 Homo sapi
c	361	18	1.2	175956	74	AC073367 AC073367 Homo sapi
c	362	18	1.2	176075	4	AC007291 AC007291 Drosophil
	363	18	1.2	176380	62	AC011371 AC011371 Homo sapi
c	364	18	1.2	176688	77	AC084696 AC084696 Mus muscu
	365	18	1.2	176733	73	AC068001 AC068001 Homo sapi
	366	18	1.2	176843	70	AC026655 AC026655 Homo sapi
	367	18	1.2	177070	65	AC018875 AC018875 Homo sapi
	368	18	1.2	177281	60	AC008494 AC008494 Homo sapi
	369	18	1.2	177386	65	AC018971 AC018971 Homo sapi
c	370	18	1.2	177562	67	AC021966 AC021966 Homo sapi
c	371	18	1.2	177703	87	AC019046 AC019046 Homo sapi
c	372	18	1.2	178242	86	AC008151 AC008151 Homo sapi
	373	18	1.2	178271	71	AC044795 AC044795 Homo sapi
c	374	18	1.2	178367	86	AC006115 AC006115 Homo sapi
c	375	18	1.2	178473	81	AL449344 AL449344 Homo sapi
c	376	18	1.2	178513	63	AC013364 AC013364 Homo sapi
c	377	18	1.2	178632	68	AC024097 AC024097 Homo sapi
	378	18	1.2	178669	64	AC016789 AC016789 Homo sapi
c	379	18	1.2	179613	81	AL390881 AL390881 Homo sapi
c	380	18	1.2	179888	92	HS1022E24 AL121581 Human DNA
	381	18	1.2	180178	69	AC025287 AC025287 Homo sapi
c	382	18	1.2	181191	76	AC080138 AC080138 Mus muscu
c	383	18	1.2	181324	90	AL160032 AL160032 Human DNA
c	384	18	1.2	181364	75	AC079226 AC079226 Homo sapi
c	385	18	1.2	182994	81	AL445924 AL445924 Homo sapi
	386	18	1.2	183147	12	AC012392 AC012392 Genomic S
c	387	18	1.2	183477	64	AC015990 AC015990 Homo sapi
	388	18	1.2	183855	61	AC009591 AC009591 Homo sapi
c	389	18	1.2	183922	60	AC005814 AC005814 Drosophil
c	390	18	1.2	184182	87	AC008952 AC008952 Homo sapi
	391	18	1.2	184430	77	AC084838 AC084838 Homo sapi
c	392	18	1.2	184824	69	AC025898 AC025898 Homo sapi
	393	18	1.2	184930	77	AC087408 AC087408 Homo sapi
c	394	18	1.2	185531	62	AC012059 AC012059 Homo sapi
	395	18	1.2	185982	74	AC073231 AC073231 Homo sapi
	396	18	1.2	186346	62	AC011596 AC011596 Homo sapi
	397	18	1.2	186558	79	AL162735 AL162735 Homo sapi
c	398	18	1.2	186797	67	AC021965 AC021965 Homo sapi
	399	18	1.2	187490	81	AL445643 AL445643 Homo sapi
	400	18	1.2	187818	82	AL589787 AL589787 Homo sapi
	401	18	1.2	187834	81	AL512605 AL512605 Homo sapi
c	402	18	1.2	188166	69	AC026124 AC026124 Homo sapi
	403	18	1.2	188168	65	AC018529 AC018529 Homo sapi
c	404	18	1.2	188346	61	AC010189 AC010189 Homo sapi
c	405	18	1.2	188351	71	AC044813 AC044813 Homo sapi
c	406	18	1.2	189035	69	AC026017 AC026017 Homo sapi
c	407	18	1.2	189351	73	AC068588 AC068588 Homo sapi
	408	18	1.2	190000	60	AC004580 AC004580 Homo sapi
	409	18	1.2	190336	73	AC068615 AC068615 Homo sapi
c	410	18	1.2	190858	86	AC007537 AC007537 Homo sapi
c	411	18	1.2	191094	77	AC087840 AC087840 Mus muscu
	412	18	1.2	191104	81	AL390962 AL390962 Homo sapi
c	413	18	1.2	191228	62	AC012356 AC012356 Homo sapi
c	414	18	1.2	191356	87	AC009455 AC009455 Homo sapi
c	415	18	1.2	191569	79	AL355474 AL355474 Homo sapi
	416	18	1.2	192003	94	AC084292 AC084292 Mus muscu
c	417	18	1.2	192015	67	AC022453 AC022453 Mus muscu
c	418	18	1.2	192440	83	CNS01DXH AL139316 Homo sapi
c	419	18	1.2	193646	60	AC008500 AC008500 Homo sapi
	420	18	1.2	194995	87	AC018696 AC018696 Homo sapi
	421	18	1.2	195290	75	AC074364 AC074364 Homo sapi
	422	18	1.2	197082	86	AC006249 AC006249 Homo sapi
c	423	18	1.2	197238	67	AC023141 AC023141 Homo sapi
c	424	18	1.2	197419	13	ATCHRIV41 AL161541 Arabidops
c	425	18	1.2	198563	13	ATCHRIV14 AL161502 Arabidops
c	426	18	1.2	199380	72	AC048379 AC048379 Homo sapi
	427	18	1.2	200000	60	AC008077 AC008077 Homo sapi
c	428	18	1.2	200165	76	AC079904 AC079904 Homo sapi
c	429	18	1.2	200252	13	ATFCA3 Z97338 Arabidopsis
	430	18	1.2	200254	71	AC034116 AC034116 Mus muscu
c	431	18	1.2	201408	88	AC090644 AC090644 Homo sapi
	432	18	1.2	201549	80	AL356130 AL356130 Homo sapi
c	433	18	1.2	203530	67	AC023048 AC023048 Mus muscu
	434	18	1.2	203861	78	AL136998 AL136998 Mus muscu
	435	18	1.2	204025	64	AC017083 AC017083 Homo sapi
c	436	18	1.2	205116	60	AC007937 AC007937 Mus muscu
c	437	18	1.2	205453	76	AC079556 AC079556 Mus muscu
c	438	18	1.2	205608	82	AL589871 AL589871 Mus muscu
c	439	18	1.2	206504	87	AC012491 AC012491 Homo sapi
	440	18	1.2	207160	64	AC016464 AC016464 Mus muscu
	441	18	1.2	207166	74	AC073703 AC073703 Mus muscu
	442	18	1.2	208660	66	AC020577 AC020577 Homo sapi
	443	18	1.2	208928	67	AC022164 AC022164 Homo sapi
c	444	18	1.2	208936	61	AC010821 AC010821 Homo sapi
	445	18	1.2	210398	87	AC010608 AC010608 Homo sapi
c	446	18	1.2	210663	65	AC019525 AC019525 Drosophil

c 447	18	1.2	211440	81	AL512642	AL512642 Homo sapi
448	18	1.2	214025	86	AC007882	AC007882 Homo sapi
449	18	1.2	214082	68	AC024540	AC024540 Homo sapi
450	18	1.2	214656	77	AC084725	AC084725 Mus muscu
c 451	18	1.2	214847	84	CNS05TCZ	AL355888 Homo sapi
c 452	18	1.2	218311	65	AC019163	AC019163 Homo sapi
c 453	18	1.2	221569	63	AC015537	AC015537 Homo sapi
c 454	18	1.2	222608	61	AC010536	AC010536 Homo sapi
455	18	1.2	227856	60	AC007908	AC007908 Homo sapi
456	18	1.2	230696	86	AC007221	AC007221 Homo sapi
c 457	18	1.2	231661	75	AC073770	AC073770 Mus muscu
c 458	18	1.2	234257	76	AC079574	AC079574 Mus muscu
459	18	1.2	234542	97	HUAC002041	AC002041 Human Chr
c 460	18	1.2	238351	81	AL445993	AL445993 Homo sapi
c 461	18	1.2	240495	5	AE003815	AE003815 Drosophil
c 462	18	1.2	245156	76	AC084021	AC084021 Mus muscu
463	18	1.2	249287	60	AC006745	AC006745 Caenorhab
464	18	1.2	251154	5	AE003656	AE003656 Drosophil
465	18	1.2	256172	60	AC005139	AC005139 Plasmodiu
466	18	1.2	257487	83	CEY64G10	Z92863 Caenorhabdi
467	18	1.2	259865	67	AC022972	AC022972 Homo sapi
468	18	1.2	265236	4	AE003626	AE003626 Drosophil
469	18	1.2	269525	4	AE003559	AE003559 Drosophil
470	18	1.2	274656	76	AC079538	AC079538 Mus muscu
c 471	18	1.2	277023	4	AE003519	AE003519 Drosophil
472	18	1.2	292943	68	AC023605	AC023605 Mus muscu
c 473	18	1.2	297900	60	AC006704	AC006704 Caenorhab
474	18	1.2	299081	60	AC006892	AC006892 Caenorhab
c 475	18	1.2	300829	4	AE003475	AE003475 Drosophil
476	18	1.2	306321	4	AE003585	AE003585 Drosophil
c 477	18	1.2	307380	77	AC087155	AC087155 Mus muscu
c 478	18	1.2	310364	4	AE003448	AE003448 Drosophil
c 479	18	1.2	310493	4	AE003480	AE003480 Drosophil
480	18	1.2	310779	60	AC005140	AC005140 Plasmodiu
c 481	18	1.2	331801	3	NMA4Z2491	AL162755 Neisseria
482	18	1.2	335283	66	AC020865	AC020865 Mus muscu
c 483	18	1.2	340000	92	HS21C080	AL163280 Homo sapi
484	18	1.2	341064	75	AC074226	AC074226 Arabidops
485	18	1.2	346820	2	AP003008	AP003008 Mesorhizo
c 486	17	1.1	147	95	RATCOL1A1	M27207 Rattus norv
487	17	1.1	170	85	AB032018	AB032018 Homo sapi
c 488	17	1.1	187	95	RNU75405	U75405 Rattus norv
c 489	17	1.1	213	95	RATCOL1A1A	M27208 Rattus norv
c 490	17	1.1	256	54	DM158G11S	Z50393 D. melanoga
491	17	1.1	303	9	A51680	A51680 Sequence 24
492	17	1.1	303	9	AX001614	AX001614 Sequence
c 493	17	1.1	313	97	HSZ95363	Z95363 H.sapiens g
c 494	17	1.1	319	97	HSZ85647	Z85647 H.sapiens B
495	17	1.1	321	54	HSA337ZH9	Z52923 H.sapiens (
496	17	1.1	366	58	AF220673	AF220673 HIV-1 AR2
497	17	1.1	399	10	AX069825	AX069825 Sequence
c 498	17	1.1	459	89	AF323474	AF323474 Leontopit
c 499	17	1.1	480	12	AF057030	AF057030 Pyrus com
500	17	1.1	488	93	HSA297074	AJ297074 Homo sapi
c 501	17	1.1	551	8	AF125119	AF125119 Nemadacty
c 502	17	1.1	566	54	HSC55E2	AL158410 H.sapiens
503	17	1.1	637	9	A13471	A13471 P.falciparu
504	17	1.1	637	10	E64858	E64858 Malaria-spe
505	17	1.1	657	5	AF030183	AF030183 Entamoeba
c 506	17	1.1	807	89	AF323475	AF323475 Homo sapi
507	17	1.1	809	71	AC035607	AC035607 Giardia i
c 508	17	1.1	813	71	AC035606	AC035606 Giardia i
c 509	17	1.1	876	94	MMU235950	AJ235950 Mus muscu
c 510	17	1.1	888	13	AF344328	AF344328 Arabidops
511	17	1.1	889	53	CNS06HLP	AL399155 T7 end of
512	17	1.1	907	59	SBA236113	AJ236113 Sericulus
c 513	17	1.1	934	8	IPU39193	U39193 Ictalurus p
514	17	1.1	942	75	AC075510	AC075510 Giardia i
c 515	17	1.1	950	77	AC086132	AC086132 Giardia i
c 516	17	1.1	959	71	AC038702	AC038702 Giardia i
517	17	1.1	973	72	AC064106	AC064106 Giardia i
518	17	1.1	977	71	AC037252	AC037252 Giardia i
519	17	1.1	986	74	AC072636	AC072636 Giardia i
c 520	17	1.1	1001	53	CNS07151	AL424475 T3 end of
c 521	17	1.1	1021	71	AC038947	AC038947 Giardia i
c 522	17	1.1	1022	53	CNS06YL3	AL421165 T3 end of
523	17	1.1	1031	7	AF107300	AF107300 Oryctolag
524	17	1.1	1055	75	AC078576	AC078576 Giardia i
525	17	1.1	1058	7	AB009313	AB009313 Oryctolag
c 526	17	1.1	1064	53	CNS079LI	AL435436 T7 end of
527	17	1.1	1212	9	AX023449	AX023449 Sequence
c 528	17	1.1	1278	1	ABIPDAGEN	X88853 A.brasilens
529	17	1.1	1333	12	AB038486	AB038486 Caulerpa
530	17	1.1	1333	12	AB038488	AB038488 Halimeda
531	17	1.1	1333	12	AB038489	AB038489 Halimeda
532	17	1.1	1395	13	AF188921	AF188921 Trichoder
533	17	1.1	1456	13	AF188918	AF188918 Hypocrea
534	17	1.1	1458	13	AF360226	AF360226 Arabidops
535	17	1.1	1472	14	ATU72503	U72503 Arabidopsis
c 536	17	1.1	1533	59	RHWNUPR	L30103 Lettuce nec
537	17	1.1	1535	89	AF330164	AF330164 Homo sapi
538	17	1.1	1666	14	D50530	D50530 Acetabulari
c 539	17	1.1	1676	8	XLIMPA1B	L36340 Xenopus lae
540	17	1.1	1708	5	AF187012	AF187012 Caenorhab
c 541	17	1.1	1722	93	HSEYA3	Y10262 H.sapiens E
c 542	17	1.1	1757	12	AF113426	AF113426 Mucor amp
543	17	1.1	1791	9	AX010004	AX010004 Sequence
544	17	1.1	1815	95	RATCOO2A	L20085 Rat cycloox
545	17	1.1	1815	95	RNU04300	U04300 Rattus norv
546	17	1.1	1827	95	RATCYLO2X	L25925 Rat cycloox
547	17	1.1	1854	93	HSHGM071	X64994 H.sapiens H
c 548	17	1.1	1862	89	AK022240	AK022240 Homo sapi
c 549	17	1.1	1921	58	AF325966	AF325966 Andes vir
550	17	1.1	1934	12	AB008097	AB008097 Arabidops
c 551	17	1.1	1965	3	MHO243902	AJ243902 Mycoplasm
c 552	17	1.1	2042	1	AF031148	AF031148 Methyloco
553	17	1.1	2068	12	AB038485	AB038485 Caulerpa
c 554	17	1.1	2085	13	AF261107	AF261107 Hordeum v
c 555	17	1.1	2147	1	AB030003	AB030003 Proteus v
556	17	1.1	2173	89	AK022312	AK022312 Homo sapi
c 557	17	1.1	2749	95	RNHLURR	X17367 Rat hepatic
558	17	1.1	2753	97	S76992	S76992 VAV2=VAV on
c 559	17	1.1	2925	6	DMMBNGEN	Z47722 D.melanogas
c 560	17	1.1	3043	3	LLPLREPAB	X60454 L.lactis pl
561	17	1.1	3258	1	AE001571	AE001571 Helicobac
c 562	17	1.1	3260	8	AB047390	AB047390 Oreochrom
c 563	17	1.1	3265	7	SSAJ3742	AJ223742 Sus scrof
564	17	1.1	3296	96	S77548	S77548 chitooligos
565	17	1.1	3324	1	AB047379	AB047379 Synechoco
566	17	1.1	3332	15	VFA277288	AJ277288 Vicia fab
567	17	1.1	3346	3	MTDNAHEMA	X83691 M.thermoaut
c 568	17	1.1	3521	9	AX057018	AX057018 Sequence
c 569	17	1.1	3541	13	ATEM1	Z11158 A.thaliana
570	17	1.1	3594	6	DDIGP2A	M77492 Dictyosteli
c 571	17	1.1	3679	8	XLESCGAM2	Y12001 X.laevis mR
572	17	1.1	3702	58	AF260511S3	AF260513 Grass car
573	17	1.1	3702	58	AF284503	AF284503 Grass car
574	17	1.1	3859	2	AF326474	AF326474 Brucella
575	17	1.1	3922	93	HSERF13	X79067 H.sapiens E
576	17	1.1	4122	97	HUM2OGDH	D10523 Human mRNA
577	17	1.1	4154	95	RRU03389	U03389 Rattus norv
578	17	1.1	4154	95	S67722	S67722 Rattus norv
579	17	1.1	4199	91	BC004964	BC004964 Homo sapi
580	17	1.1	4229	89	AK024501	AK024501 Homo sapi
581	17	1.1	4404	94	AF233596	AF233596 Rattus no
582	17	1.1	4412	3	LMMETYX	Y10744 L.meyeri me
583	17	1.1	5198	15	SCYGR288W	Z73073 S.cerevisia
584	17	1.1	5214	85	AB046794	AB046794 Homo sapi
585	17	1.1	5482	97	HUM2OD5	D32060 Human gene
c 586	17	1.1	5489	1	AF094574	AF094574 Haemophil
587	17	1.1	5499	1	AF039082	AF039082 Lactococc
588	17	1.1	5715	3	STU51927	U51927 Salmonella
c 589	17	1.1	5746	9	A51694	A51694 Sequence 38
c 590	17	1.1	5746	9	AX001628	AX001628 Sequence
c 591	17	1.1	5826	63	AC014091	AC014091 Drosophil
c 592	17	1.1	7869	14	LLVTRNAG	Z49255 L.luteus tR

	593	17	1.1	8457	3	STAJ4892 AJ224892 Salmonell
	594	17	1.1	8457	9	AX023428 AX023428 Sequence
	595	17	1.1	8527	91	CJINTERFG X64659 C.jacchus i
c	596	17	1.1	9223	97	HUMMDCDNA D31872 Human metal
c	597	17	1.1	9278	10	I25849 I25849 Sequence 9
c	598	17	1.1	9278	10	I43486 I43486 Sequence 9
c	599	17	1.1	9278	45	E10494 E10494 gDNA encodi
	600	17	1.1	10126	89	AF119817 AF119817 Homo sapi
	601	17	1.1	10141	89	AF119818 AF119818 Homo sapi
	602	17	1.1	10185	3	U32754 U32754 Haemophilus
	603	17	1.1	10244	1	AE002312 AE002312 Chlamydia
	604	17	1.1	10384	3	U32692 U32692 Haemophilus
c	605	17	1.1	10681	94	MMOR23 X92969 M.musculus
c	606	17	1.1	11075	1	AE004269 AE004269 Vibrio ch
c	607	17	1.1	11272	94	AF022247 AF022247 Rattus no
	608	17	1.1	12165	1	AE004328 AE004328 Vibrio ch
	609	17	1.1	12940	4	AE001386 AE001386 Plasmodiu
c	610	17	1.1	13323	96	PFA245435 AJ245435 Plasmodiu
	611	17	1.1	13736	64	AC017291 AC017291 Drosophil
c	612	17	1.1	13975	93	HSCGMPPDE X94354 H.sapiens D
c	613	17	1.1	14999	1	AE004910 AE004910 Pseudomon
c	614	17	1.1	15211	64	AC017412 AC017412 Drosophil
c	615	17	1.1	15288	6	DROFASI M32311 D.melanogas
c	616	17	1.1	15984	5	AF217843 AF217843 Penaeus m
c	617	17	1.1	17446	6	CEY111B2C AL132906 Caenorhab
	618	17	1.1	17742	63	AC014824 AC014824 Drosophil
c	619	17	1.1	17758	6	CELT21F4 U56959 Caenorhabdi
c	620	17	1.1	18274	58	AB025319 AB025319 Yaba monk
c	621	17	1.1	20803	6	CELF57H12 AF045644 Caenorhab
c	622	17	1.1	21332	63	AC015231 AC015231 Drosophil
	623	17	1.1	23703	94	AF220365 AF220365 Mus muscu
	624	17	1.1	23715	65	AC018130 AC018130 Drosophil
	625	17	1.1	25949	3	MTCY1A10 Z95387 Mycobacteri
c	626	17	1.1	26263	58	AF074946 AF074946 Hemorrhag
c	627	17	1.1	26270	9	AX009977 AX009977 Sequence
c	628	17	1.1	27736	90	AL391426 AL391426 Human DNA
	629	17	1.1	29303	6	CER08H2 Z81575 Caenorhabdi
	630	17	1.1	29839	93	HSL33C6 Z68282 Human DNA s
	631	17	1.1	30590	3	SC4G2 AL031371 Streptomy
	632	17	1.1	30769	87	AC009037 AC009037 Homo sapi
c	633	17	1.1	31226	6	CELZK622 U39998 Caenorhabdi
	634	17	1.1	31661	6	CELM60 U39995 Caenorhabdi
c	635	17	1.1	32481	6	CELF40H6 U21310 Caenorhabdi
	636	17	1.1	33834	64	AC017478 AC017478 Drosophil
	637	17	1.1	34013	97	HSU73634 U73634 Human chrom
	638	17	1.1	34037	6	CELF43C9 U40427 Caenorhabdi
	639	17	1.1	34305	6	CEW02B12 Z66521 Caenorhabdi
	640	17	1.1	34569	65	AC020276 AC020276 Drosophil
	641	17	1.1	35031	5	AF052833 AF052833 Trypanoso
c	642	17	1.1	35064	5	AF163772 AF163772 Leishmani
c	643	17	1.1	35331	88	AF068624 AF068624 Homo sapi
c	644	17	1.1	35809	89	AF153832S5 AF153836 Homo sapi
	645	17	1.1	36255	64	AC016208 AC016208 Homo sapi
	646	17	1.1	36527	77	AC087396 AC087396 Homo sapi
	647	17	1.1	36534	65	AC018212 AC018212 Drosophil
c	648	17	1.1	36588	64	AC017584 AC017584 Drosophil
c	649	17	1.1	36998	6	CELC54A12 U28733 Caenorhabdi
	650	17	1.1	37777	15	SPBC215 AL033534 S.pombe c
c	651	17	1.1	37881	6	CELK09H9 AF043700 Caenorhab
c	652	17	1.1	38700	6	CET09A5 Z36753 Caenorhabdi
c	653	17	1.1	38779	6	CET24B8 Z68338 Caenorhabdi
	654	17	1.1	38946	6	LMFL1232 AL354512 Leishmani
	655	17	1.1	38984	65	AC020415 AC020415 Drosophil
	656	17	1.1	39021	88	AF052041 AF052041 Homo sapi
	657	17	1.1	39038	97	HSU78308 U78308 Human olfac
	658	17	1.1	39437	61	AC010701 AC010701 Drosophil
c	659	17	1.1	39613	6	CELZK328 U50193 Caenorhabdi
c	660	17	1.1	39702	85	AC002111 AC002111 Genomic s
c	661	17	1.1	39784	5	CEC14A4 Z49909 Caenorhabdi
	662	17	1.1	39904	91	AL590003 AL590003 Human DNA
c	663	17	1.1	39963	85	AC000044 AC000044 Homo sapi
c	664	17	1.1	40226	63	AC013338 AC013338 Homo sapi
c	665	17	1.1	40583	86	AC005621 AC005621 Homo sapi
	666	17	1.1	42293	68	AC023531 AC023531 Homo sapi
c	667	17	1.1	43740	85	AC000034 AC000034 Homo sapi
	668	17	1.1	44036	6	CELT28A11 U80027 Caenorhabdi
c	669	17	1.1	45005	85	AC004511 AC004511 Homo sapi
c	670	17	1.1	47745	83	CET09D4 Z81590 Caenorhabdi
	671	17	1.1	48060	90	AL139798 AL139798 Human DNA
c	672	17	1.1	48128	9	AX059474 AX059474 Sequence
	673	17	1.1	51461	60	AC008607 AC008607 Homo sapi
	674	17	1.1	51471	87	AC008724 AC008724 Homo sapi
c	675	17	1.1	51979	65	AC018743_3 Continuation (4 of
c	676	17	1.1	56295	85	AC004011 AC004011 Human BAC
	677	17	1.1	56934	82	AP000582 AP000582 Homo sapi
c	678	17	1.1	57264	77	AC087477 AC087477 Homo sapi
	679	17	1.1	57882	74	AC069347 AC069347 Homo sapi
c	680	17	1.1	58752	60	AC008295 AC008295 Drosophil
	681	17	1.1	59015	73	AC067811 AC067811 Homo sapi
c	682	17	1.1	59015	73	AC067811 AC067811 Homo sapi
	683	17	1.1	59412	14	NCB7F21 AL389901 Neurospor
	684	17	1.1	59865	63	AC014422 AC014422 Drosophil
c	685	17	1.1	59865	63	AC014422 AC014422 Drosophil
c	686	17	1.1	60000	12	AF049236 AF049236 Arabidops
	687	17	1.1	60116	78	AC090933 AC090933 Homo sapi
	688	17	1.1	60129	86	AC006325 AC006325 Homo sapi
c	689	17	1.1	60602	65	AC018076 AC018076 Drosophil
c	690	17	1.1	61024	77	AC087302 AC087302 Homo sapi
	691	17	1.1	62040	76	AC080123 AC080123 Homo sapi
	692	17	1.1	62043	65	AC018986 AC018986 Mus muscu
	693	17	1.1	63646	93	HSJ256K2A Z72004 Human DNA s
	694	17	1.1	63710	76	AC084183 AC084183 Homo sapi
c	695	17	1.1	64433	68	AC024001 AC024001 Homo sapi
	696	17	1.1	64507	77	AC087712 AC087712 Homo sapi
c	697	17	1.1	64772	78	AC090588 AC090588 Homo sapi
	698	17	1.1	65354	64	AC016807 AC016807 Homo sapi
c	699	17	1.1	65467	64	AC017402 AC017402 Drosophil
	700	17	1.1	65554	62	AC011204 AC011204 Homo sapi
	701	17	1.1	65691	96	PFMAL3P1 Z97348 Plasmodium
	702	17	1.1	65792	10	AX067456 AX067456 Sequence
c	703	17	1.1	66752	78	AC090366 AC090366 Homo sapi
c	704	17	1.1	67870	76	AC083787 AC083787 Homo sapi
c	705	17	1.1	68050	75	AC079074 AC079074 Homo sapi
	706	17	1.1	68487	65	AC017782 AC017782 Drosophil
	707	17	1.1	69598	68	AC023647 AC023647 Homo sapi
c	708	17	1.1	69978	76	AC080107 AC080107 Homo sapi
	709	17	1.1	70548	93	HSAF003625 AF003625 Homo sapi
c	710	17	1.1	70592	83	AP002017 AP002017 Homo sapi
c	711	17	1.1	70610	64	AC017561 AC017561 Drosophil
c	712	17	1.1	71287	63	AC014345 AC014345 Drosophil
	713	17	1.1	72615	73	AC068678 AC068678 Homo sapi
c	714	17	1.1	72907	64	AC016010 AC016010 Homo sapi
	715	17	1.1	73029	78	AC090560 AC090560 Homo sapi
	716	17	1.1	73339	70	AC027246 AC027246 Homo sapi
	717	17	1.1	74795	78	AC091249 AC091249 Mus muscu
	718	17	1.1	75021	71	AC040889 AC040889 Homo sapi
	719	17	1.1	75335	92	HS107715 AL034394 Human DNA
c	720	17	1.1	75424	62	AC011867 AC011867 Homo sapi
c	721	17	1.1	75799	14	F26G16 AC009917 Arabidops
c	722	17	1.1	76619	4	AC004722 AC004722 Drosophil
	723	17	1.1	76660	69	AC025153 AC025153 Homo sapi
c	724	17	1.1	77447	6	L31790 L31790 Drosophila
	725	17	1.1	77758	12	AB016885 AB016885 Arabidops
	726	17	1.1	79562	93	HSJ321O10 AL109752 Human DNA
	727	17	1.1	79722	78	AL139346 AL139346 Homo sapi
	728	17	1.1	79779	93	HSJ710H13 AL121712 Human DNA
c	729	17	1.1	81435	83	CEY73F8_3 Continuation (4 of
c	730	17	1.1	82454	12	AC012193 AC012193 Arabidops
	731	17	1.1	84371	83	AP001963 AP001963 Homo sapi
	732	17	1.1	85109	14	ATT2J13 AL132967 Arabidops
c	733	17	1.1	85654	93	HSP373C6 AL022393 Homo sapi
	734	17	1.1	86005	79	AL355475 AL355475 Homo sapi
c	735	17	1.1	86050	90	AL359647 AL359647 Human DNA
c	736	17	1.1	86152	86	AC008585 AC008585 Homo sapi
c	737	17	1.1	86358	12	AC069326 AC069326 Genomic S
	738	17	1.1	86574	14	ATT15N1 AL163792 Arabidops

	739	17	1.1	87233	61	AC010613 AC010613 Homo sapi
c	740	17	1.1	87394	12	AC005882 AC005882 Arabidops
	741	17	1.1	87578	61	AC010474 AC010474 Homo sapi
c	742	17	1.1	87799	12	AB024036 AB024036 Arabidops
	743	17	1.1	88165	63	AC014754 AC014754 Drosophil
	744	17	1.1	89181	63	AC012641 AC012641 Homo sapi
c	745	17	1.1	89670	65	AC017901 AC017901 Drosophil
c	746	17	1.1	90052	90	AL139429 AL139429 Human DNA
	747	17	1.1	90109	60	AC002557 AC002557 Homo sapi
	748	17	1.1	90220	93	HSJ734P14 AL049650 Human DNA
	749	17	1.1	91040	64	AC017292 AC017292 Drosophil
	750	17	1.1	91202	60	AC006186 AC006186 Homo sapi
c	751	17	1.1	91299	12	AC083859 AC083859 Arabidops
c	752	17	1.1	91413	12	AC022456 AC022456 Arabidops
c	753	17	1.1	91875	66	AC020926 AC020926 Homo sapi
c	754	17	1.1	92179	90	AL390965 AL390965 Human DNA
c	755	17	1.1	92620	12	AB026636 AB026636 Arabidops
c	756	17	1.1	92807	92	HS1137F22 AL034421 Human DNA
	757	17	1.1	92870	88	AC022425 AC022425 Homo sapi
c	758	17	1.1	92918	65	AC020050 AC020050 Drosophil
	759	17	1.1	93553	61	AC010593 AC010593 Homo sapi
	760	17	1.1	93829	4	AC007122 AC007122 Drosophil
c	761	17	1.1	94461	12	AC004684 AC004684 Arabidops
c	762	17	1.1	94695	13	ATF23E13 AL022141 Arabidops
	763	17	1.1	95780	60	AC008416 AC008416 Homo sapi
	764	17	1.1	95808	78	AF214636 AF214636 Homo sapi
	765	17	1.1	96270	93	HSDJ833K8 AL096704 Human DNA
c	766	17	1.1	96729	67	AC023171 AC023171 Homo sapi
	767	17	1.1	97126	78	AF205590 AF205590 Homo sapi
c	768	17	1.1	97234	89	AL137859 AL137859 Human DNA
	769	17	1.1	97898	90	AL353613 AL353613 Human DNA
	770	17	1.1	98567	63	AC013825 AC013825 Homo sapi
c	771	17	1.1	98573	88	AF095725 AF095725 Homo sapi
c	772	17	1.1	98708	73	AC068510 AC068510 Homo sapi
	773	17	1.1	99011	60	AC004937 AC004937 Homo sapi
	774	17	1.1	99271	92	HS808P6 AL034411 Human DNA
c	775	17	1.1	99587	87	AC010498 AC010498 Homo sapi
	776	17	1.1	99850	89	AL136081 AL136081 Human DNA
	777	17	1.1	100269	13	ATF18O22 AL163817 Arabidops
	778	17	1.1	100375	92	HS193B12 Z98744 Human DNA s
	779	17	1.1	100452	85	AC004499 AC004499 Homo sapi
c	780	17	1.1	100496	92	HS970D1 AL021069 Human DNA
c	781	17	1.1	100706	76	AC079338 AC079338 Homo sapi
c	782	17	1.1	101458	14	F5O24 AF296825 Arabidops
c	783	17	1.1	102473	60	AC002320 AC002320 Homo sapi
	784	17	1.1	103146	92	HS187N21 Z98036 Human DNA s
	785	17	1.1	103911	97	HUAC002565 AC002565 Human Chr
	786	17	1.1	104278	4	AC005269 AC005269 Drosophil
	787	17	1.1	104377	78	AF235105 AF235105 Homo sapi
c	788	17	1.1	105160	61	AC010571 AC010571 Drosophil
c	789	17	1.1	105733	75	AC074136 AC074136 Homo sapi
	790	17	1.1	105787	87	AC010458 AC010458 Homo sapi
	791	17	1.1	106039	85	AC004159 AC004159 Homo sapi
	792	17	1.1	106073	89	AL136146 AL136146 Human DNA
	793	17	1.1	106482	94	AC007585 AC007585 Mus muscu
c	794	17	1.1	106702	13	ATF7K2 AL033545 Arabidops
	795	17	1.1	107655	77	AC084840 AC084840 Homo sapi
c	796	17	1.1	108131	90	AL139321 AL139321 Human DNA
c	797	17	1.1	108509	76	AC079391 AC079391 Homo sapi
c	798	17	1.1	108561	60	AC007885 AC007885 Drosophil
c	799	17	1.1	108726	79	AL355866 AL355866 Homo sapi
	800	17	1.1	109018	93	HSJ1129A6 AL096763 Human DNA
c	801	17	1.1	109047	84	HSDJ19F5 AL078592 Homo sapi
	802	17	1.1	109097	63	AC013377 AC013377 Homo sapi
c	803	17	1.1	109230	80	AL356362 AL356362 Homo sapi
c	804	17	1.1	109740	65	AC020430 AC020430 Drosophil
	805	17	1.1	110000	60	AC007194_0 AC007194 Homo sapi
c	806	17	1.1	110000	62	AC011329_1 Continuation (2 of
	807	17	1.1	110000	62	AC011600_2 Continuation (3 of
	808	17	1.1	110000	68	AC024562_1 Continuation (2 of
c	809	17	1.1	110000	83	CEY111B2_0 Z98857 Caenorhabdi
c	810	17	1.1	110000	83	CEY73F8_2 Continuation (3 of
c	811	17	1.1	110000	84	TBBCHR1A_06 Continuation (7 of
	812	17	1.1	110000	85	AC005079_0 AC005079 Homo sapi
c	813	17	1.1	110565	12	AC009243 AC009243 Genomic s
c	814	17	1.1	110638	81	AL391725 AL391725 Homo sapi
c	815	17	1.1	111123	72	AC066604 AC066604 Homo sapi
	816	17	1.1	111799	93	HSJ850O15 AL049637 Human DNA
c	817	17	1.1	111994	92	HS296K21 Z83821 Human DNA s
	818	17	1.1	112864	68	AC024127 AC024127 Homo sapi
	819	17	1.1	113075	91	AL512363 AL512363 Human DNA
	820	17	1.1	113146	90	AL354856 AL354856 Human DNA
c	821	17	1.1	113320	61	AC010662 AC010662 Drosophil
	822	17	1.1	113539	90	AL353681 AL353681 Human DNA
c	823	17	1.1	113970	14	ATT6G15 AL049656 Arabidops
c	824	17	1.1	114173	92	HS884M20 AL020991 Homo sapi
c	825	17	1.1	114839	78	AF165140 AF165140 Homo sapi
c	826	17	1.1	115405	78	AL138793 AL138793 Homo sapi
c	827	17	1.1	115866	80	AL358944 AL358944 Homo sapi
	828	17	1.1	116019	85	AC004629 AC004629 Homo sapi
c	829	17	1.1	116341	71	AC044802 AC044802 Homo sapi
c	830	17	1.1	116397	90	AL157902 AL157902 Human DNA
	831	17	1.1	117273	63	AC015138 AC015138 Drosophil
	832	17	1.1	117522	91	AL512505 AL512505 Human DNA
c	833	17	1.1	118423	78	AF252825 AF252825 Homo sapi
c	834	17	1.1	118923	60	AC008922 AC008922 Homo sapi
	835	17	1.1	119407	13	ATF5K20 AL132960 Arabidops
	836	17	1.1	119420	92	HS997K18 AL021406 Human DNA
	837	17	1.1	120181	63	AC014784 AC014784 Drosophil
	838	17	1.1	120391	64	AC017146 AC017146 Drosophil
c	839	17	1.1	120608	92	HS469K11 AL449209 Homo sapi
	840	17	1.1	120998	93	HSDJ726A1 AL109842 Human DNA
	841	17	1.1	121029	90	AL391726 AL391726 Human DNA
	842	17	1.1	121414	64	AC017234 AC017234 Drosophil
c	843	17	1.1	122224	70	AC026397 AC026397 Homo sapi
	844	17	1.1	122764	93	HSJ831D17 AL109984 Human DNA
	845	17	1.1	122914	94	AC069451 AC069451 Mus muscu
	846	17	1.1	123160	76	AC083773 AC083773 Homo sapi
c	847	17	1.1	123323	91	AP001816 AP001816 Homo sapi
c	848	17	1.1	123608	84	HS330O12 AL031731 Homo sapi
	849	17	1.1	123782	93	HSDJ756N5 AL132825 Human DNA
	850	17	1.1	124010	92	HS340G1 Z84719 Human DNA s
	851	17	1.1	124131	64	AC016335 AC016335 Homo sapi
c	852	17	1.1	124321	92	HS417G15 AL009174 Human DNA
	853	17	1.1	124347	87	AC010072 AC010072 Homo sapi
	854	17	1.1	124645	97	HUAC004638 AC004638 Homo sapi
c	855	17	1.1	124672	62	AC011403 AC011403 Homo sapi
c	856	17	1.1	125201	90	AL391097 AL391097 Human DNA
c	857	17	1.1	125760	76	AC080122 AC080122 Homo sapi
c	858	17	1.1	125768	70	AC026451 AC026451 Homo sapi
	859	17	1.1	125937	92	HS69M21 AL031735 Human DNA
	860	17	1.1	125977	84	CNS05TC5 AL355053 Homo sapi
	861	17	1.1	126033	88	AC027312 AC027312 Homo sapi
	862	17	1.1	126045	91	HS1010E17 AL121826 Human DNA
c	863	17	1.1	126147	82	AL590030 AL590030 Homo sapi
c	864	17	1.1	126327	85	AC005252 AC005252 Homo sapi
c	865	17	1.1	126482	89	AL137220 AL137220 Human DNA
c	866	17	1.1	126615	85	AC005180 AC005180 Homo sapi
c	867	17	1.1	127035	4	AC006574 AC006574 Drosophil
	868	17	1.1	127366	6	DROSADH10 AE003416 Drosophil
	869	17	1.1	127759	75	AC079021 AC079021 Oryza sat
	870	17	1.1	127794	89	AL138688 AL138688 Human DNA
c	871	17	1.1	128133	89	AL138787 AL138787 Human DNA
c	872	17	1.1	128320	89	AL138878 AL138878 Human DNA
	873	17	1.1	128529	69	AC025758 AC025758 Homo sapi
	874	17	1.1	128590	61	AC009557 AC009557 Homo sapi
c	875	17	1.1	128590	61	AC009557 AC009557 Homo sapi
c	876	17	1.1	128702	87	AC008872 AC008872 Homo sapi
	877	17	1.1	128877	70	AC026399 AC026399 Homo sapi
c	878	17	1.1	128959	86	AC006384 AC006384 Homo sapi
c	879	17	1.1	129014	85	AC004916 AC004916 Homo sapi
c	880	17	1.1	129293	87	AC010638 AC010638 Homo sapi
	881	17	1.1	129430	87	AC008905 AC008905 Homo sapi
	882	17	1.1	129756	92	HS430A16 AL022725 Human DNA
c	883	17	1.1	129757	13	ATF2K15 AL132956 Arabidops
	884	17	1.1	129955	61	AC010056 AC010056 Drosophil

885	17	1.1	129955	62	AC011909	AC011909 Drosophil
886	17	1.1	130526	82	AP000715	AP000715 Homo sapi
c 887	17	1.1	130583	60	AC007420	AC007420 Drosophil
c 888	17	1.1	130797	77	AC087902	AC087902 Mus muscu
889	17	1.1	130886	87	AC010477	AC010477 Homo sapi
c 890	17	1.1	130965	90	AL162457	AL162457 Human DNA
891	17	1.1	131196	71	AC037473	AC037473 Homo sapi
c 892	17	1.1	131277	76	AC080066	AC080066 Homo sapi
893	17	1.1	131552	89	AL137063	AL137063 Human DNA
c 894	17	1.1	131948	71	AC034183	AC034183 Homo sapi
c 895	17	1.1	131980	71	AC034267	AC034267 Homo sapi
c 896	17	1.1	132396	93	HSJ673D20	AL109809 Human DNA
c 897	17	1.1	132493	85	AC004583	AC004583 Human Chr
898	17	1.1	132727	88	AC024121	AC024121 Homo sapi
c 899	17	1.1	133391	93	HSDJ315G1	AL121601 Human DNA
c 900	17	1.1	133574	92	HS20N2	AL031320 Human DNA
c 901	17	1.1	134412	87	AC010428	AC010428 Homo sapi
902	17	1.1	134472	85	AC005186	AC005186 Human Chr
903	17	1.1	134771	61	AC010961	AC010961 Homo sapi
c 904	17	1.1	134794	82	AL589992	AL589992 Homo sapi
c 905	17	1.1	135065	88	AC026691	AC026691 Homo sapi
c 906	17	1.1	135123	82	AP001822	AP001822 Homo sapi
907	17	1.1	135244	68	AC024572	AC024572 Homo sapi
c 908	17	1.1	135252	81	AL512423	AL512423 Homo sapi
909	17	1.1	135391	92	HS805C22	AL035667 Human DNA
910	17	1.1	135672	93	HSAC000113	AC000113 Human PAC
c 911	17	1.1	135751	60	AC008586	AC008586 Homo sapi
c 912	17	1.1	135855	92	HS135E14	AJ010598 Homo sapi
913	17	1.1	135889	77	AC084817	AC084817 Oryza sat
914	17	1.1	136004	82	AL590126	AL590126 Homo sapi
c 915	17	1.1	136040	67	AC022778	AC022778 Homo sapi
916	17	1.1	136559	67	AC022647	AC022647 Homo sapi
917	17	1.1	136579	92	HS998G20	AL035246 Human DNA
c 918	17	1.1	137241	68	AC024323	AC024323 Homo sapi
919	17	1.1	137569	76	AC084093	AC084093 Homo sapi
920	17	1.1	137648	68	AC024316	AC024316 Homo sapi
921	17	1.1	137785	92	HS345P10	Z82201 Human DNA s
c 922	17	1.1	137895	89	AL137794	AL137794 Human DNA
c 923	17	1.1	137993	83	AP003201	AP003201 Oryza sat
924	17	1.1	138383	68	AC024502	AC024502 Homo sapi
925	17	1.1	138603	71	AC034182	AC034182 Homo sapi
926	17	1.1	138698	60	AC008454	AC008454 Homo sapi
c 927	17	1.1	139030	89	AL136460	AL136460 Human DNA
928	17	1.1	139152	13	AP002525	AP002525 Oryza sat
929	17	1.1	139714	68	AC024322	AC024322 Homo sapi
c 930	17	1.1	139714	68	AC024322	AC024322 Homo sapi
931	17	1.1	139934	97	HUB384D8	U62317 Homo sapien
c 932	17	1.1	140091	90	AL445684	AL445684 Human DNA
c 933	17	1.1	140210	60	AC002993	AC002993 Homo sapi
934	17	1.1	140426	80	AL365216	AL365216 Homo sapi
935	17	1.1	140488	71	AC034288	AC034288 Homo sapi
c 936	17	1.1	141012	66	AC021769	AC021769 Homo sapi
937	17	1.1	141287	72	AC063978	AC063978 Homo sapi
938	17	1.1	141710	62	AC012348	AC012348 Homo sapi
939	17	1.1	141984	14	F9L1	AC007591 Arabidops
940	17	1.1	142021	78	AC090553	AC090553 Sus scrof
941	17	1.1	142329	72	AC046155	AC046155 Homo sapi
942	17	1.1	143236	73	AC069054	AC069054 Homo sapi
c 943	17	1.1	143391	86	AC005509	AC005509 Homo sapi
944	17	1.1	143393	63	AC015222	AC015222 Drosophil
945	17	1.1	143710	13	AP002899	AP002899 Oryza sat
c 946	17	1.1	144013	69	AC025059	AC025059 Homo sapi
c 947	17	1.1	144046	73	AC068334	AC068334 Homo sapi
948	17	1.1	144198	65	AC018929	AC018929 Oryza sat
949	17	1.1	144392	92	HS67K17	AL023584 Human DNA
950	17	1.1	144546	85	AC004870	AC004870 Homo sapi
951	17	1.1	144695	62	AC012438	AC012438 Homo sapi
952	17	1.1	144734	82	AP001839	AP001839 Homo sapi
c 953	17	1.1	144759	92	HS352A20	AL021939 Homo sapi
954	17	1.1	145101	66	AC021775	AC021775 Homo sapi
c 955	17	1.1	145189	67	AC021968	AC021968 Homo sapi
956	17	1.1	145409	66	AC021260	AC021260 Homo sapi
957	17	1.1	145442	85	AC004859	AC004859 Homo sapi
c 958	17	1.1	145695	63	AC015690	AC015690 Homo sapi
c 959	17	1.1	145891	73	AC067898	AC067898 Homo sapi
960	17	1.1	146010	87	AC009289	AC009289 Homo sapi
c 961	17	1.1	146150	73	AC068333	AC068333 Homo sapi
c 962	17	1.1	146258	68	AC023479	AC023479 Homo sapi
963	17	1.1	146438	93	HSBA191L9	AL117329 Human DNA
964	17	1.1	146602	84	CNS05TD5	AL356015 Homo sapi
c 965	17	1.1	146638	83	AP003252	AP003252 Oryza sat
966	17	1.1	146670	91	AP000949	AP000949 Homo sapi
967	17	1.1	146862	80	AL390064	AL390064 Homo sapi
c 968	17	1.1	146991	87	AC012603	AC012603 Homo sapi
c 969	17	1.1	147021	86	AC006501	AC006501 Homo sapi
970	17	1.1	147095	72	AC055752	AC055752 Homo sapi
971	17	1.1	147102	85	AC002476	AC002476 Human PAC
c 972	17	1.1	147151	71	AC037458	AC037458 Homo sapi
c 973	17	1.1	147250	62	AC011778	AC011778 Homo sapi
c 974	17	1.1	147859	62	AC011079	AC011079 Homo sapi
c 975	17	1.1	147945	62	AC012244	AC012244 Homo sapi
976	17	1.1	148083	89	AF111170	AF111170 Homo sapi
977	17	1.1	148204	74	AC069400	AC069400 Homo sapi
c 978	17	1.1	148410	87	AC010654	AC010654 Homo sapi
c 979	17	1.1	148429	82	AL590375	AL590375 Homo sapi
980	17	1.1	148474	63	AC013316	AC013316 Homo sapi
981	17	1.1	148517	62	AC011147	AC011147 Homo sapi
982	17	1.1	148896	80	AL365439	AL365439 Homo sapi
c 983	17	1.1	148965	72	AC060784	AC060784 Homo sapi
c 984	17	1.1	149006	88	AC025812	AC025812 Homo sapi
c 985	17	1.1	149075	76	AC084209	AC084209 Homo sapi
c 986	17	1.1	149180	63	AC015525	AC015525 Homo sapi
c 987	17	1.1	149202	74	AC073467	AC073467 Homo sapi
c 988	17	1.1	149230	65	AC018790	AC018790 Homo sapi
989	17	1.1	149354	80	AL358032	AL358032 Homo sapi
c 990	17	1.1	149461	65	AC019282	AC019282 Homo sapi
c 991	17	1.1	149498	83	AP003196	AP003196 Oryza sat
c 992	17	1.1	149649	71	AC034218	AC034218 Homo sapi
993	17	1.1	149764	85	AC002076	AC002076 Human BAC
c 994	17	1.1	149924	74	AC073067	AC073067 Homo sapi
c 995	17	1.1	149969	68	AC024180	AC024180 Homo sapi
c 996	17	1.1	150020	74	AC073151	AC073151 Mus muscu
997	17	1.1	150150	63	AC015670	AC015670 Homo sapi
c 998	17	1.1	150173	81	AL583824	AL583824 Homo sapi
999	17	1.1	150224	93	HSJ1059L7	AL121913 Human DNA
c1000	17	1.1	150331	64	AC016587	AC016587 Homo sapi

ALIGNMENTS

```
RESULT  1
AC023733
LOCUS       AC023733    46913 bp    DNA              HTG       18-FEB-2000
DEFINITION  Drosophila melanogaster clone RP98-16J21, *** SEQUENCING IN
            PROGRESS ***, 41 unordered pieces.
ACCESSION   AC023733
VERSION     AC023733.2  GI:6996821
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 46913)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
            Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
```

```
            Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
            Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
            Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
            Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
            Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M.,
            Watlington,S., Weinstock,G., Weinstock,I.R., Williamson,A.,
            Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 46913)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-FEB-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Feb 18, 2000 this sequence version replaced gi:6984307.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: DRJW
                Center clone name: RP98-16J21
            -------------- Summary Statistics
                Sequencing vector: M13; L08821
                Chemistry: Dye-terminator Big Dye: 99% of reads
                Assembly program: Phrap; version 0.980611
                Consensus quality: 15493 bases at least Q40
                Consensus quality: 25427 bases at least Q30
                Consensus quality: 31480 bases at least Q20
                Estimated insert size: 29542; sum-of-contigs estimation
                Quality coverage: 0.5x in Q20 bases; sum-of-contigs estimation
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 41 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      755: contig of 755 bp in length
            *      756      775: gap of unknown length
            *      776     1598: contig of 823 bp in length
            *     1599     1618: gap of unknown length
            *     1619     3169: contig of 1551 bp in length
            *     3170     3189: gap of unknown length
            *     3190     3996: contig of 807 bp in length
            *     3997     4016: gap of unknown length
            *     4017     4859: contig of 843 bp in length
            *     4860     4879: gap of unknown length
            *     4880     5855: contig of 976 bp in length
            *     5856     5875: gap of unknown length
            *     5876     6968: contig of 1093 bp in length
            *     6969     6988: gap of unknown length
            *     6989     7947: contig of 959 bp in length
            *     7948     7967: gap of unknown length
            *     7968     8736: contig of 769 bp in length
            *     8737     8756: gap of unknown length
            *     8757     9500: contig of 744 bp in length
            *     9501     9520: gap of unknown length
            *     9521    10503: contig of 983 bp in length
            *    10504    10523: gap of unknown length
            *    10524    11456: contig of 933 bp in length
            *    11457    11476: gap of unknown length
            *    11477    12258: contig of 782 bp in length
            *    12259    12278: gap of unknown length
            *    12279    13525: contig of 1247 bp in length
            *    13526    13545: gap of unknown length
            *    13546    14301: contig of 756 bp in length
            *    14302    14321: gap of unknown length
            *    14322    15585: contig of 1264 bp in length
            *    15586    15605: gap of unknown length
            *    15606    16413: contig of 808 bp in length
            *    16414    16433: gap of unknown length
            *    16434    17352: contig of 919 bp in length
            *    17353    17372: gap of unknown length
            *    17373    18409: contig of 1037 bp in length
            *    18410    18429: gap of unknown length
            *    18430    19504: contig of 1075 bp in length
            *    19505    19524: gap of unknown length
            *    19525    20625: contig of 1101 bp in length
            *    20626    20645: gap of unknown length
            *    20646    21019: contig of 374 bp in length
            *    21020    21039: gap of unknown length
            *    21040    21850: contig of 811 bp in length
            *    21851    21870: gap of unknown length
            *    21871    22700: contig of 830 bp in length
            *    22701    22720: gap of unknown length
            *    22721    23720: contig of 1000 bp in length
            *    23721    23740: gap of unknown length
            *    23741    25116: contig of 1376 bp in length
            *    25117    25136: gap of unknown length
            *    25137    26853: contig of 1717 bp in length
            *    26854    26873: gap of unknown length
            *    26874    27573: contig of 700 bp in length
            *    27574    27593: gap of unknown length
            *    27594    28820: contig of 1227 bp in length
            *    28821    28840: gap of unknown length
            *    28841    29818: contig of 978 bp in length
            *    29819    29838: gap of unknown length
            *    29839    31525: contig of 1687 bp in length
            *    31526    31545: gap of unknown length
            *    31546    32670: contig of 1125 bp in length
            *    32671    32690: gap of unknown length
            *    32691    33781: contig of 1091 bp in length
            *    33782    33801: gap of unknown length
            *    33802    35134: contig of 1333 bp in length
            *    35135    35154: gap of unknown length
            *    35155    37141: contig of 1987 bp in length
            *    37142    37161: gap of unknown length
            *    37162    38615: contig of 1454 bp in length
            *    38616    38635: gap of unknown length
            *    38636    40858: contig of 2223 bp in length
            *    40859    40878: gap of unknown length
            *    40879    42886: contig of 2008 bp in length
            *    42887    42906: gap of unknown length
            *    42907    44311: contig of 1405 bp in length
            *    44312    44331: gap of unknown length
            *    44332    45841: contig of 1510 bp in length
            *    45842    45861: gap of unknown length
            *    45862    46913: contig of 1052 bp in length.
FEATURES             Location/Qualifiers
     source          1. .46913
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RP98-16J21"
BASE COUNT    12519 a  10639 c  10573 g  12323 t    859 others
ORIGIN


  Query Match           43.8%;  Score 673;  DB 68;  Length 46913;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  673;  Conservative    0;  Mismatches  0;  Indels    0;  Gaps    0;

Qy       1 atgaggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   29953 ATGAGGCTTCCTTGGGTGATTTTCTGTACAGTTCTGCTTTTAATTTTTACGAACAACTCA 30012

Qy      61 aatgcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaagg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   30013 AATGCAGATATTCCCGGGTGCAACTACTACGACACGGTTGATATCTCATACATTGAAAGG 30072

Qy     121 caaaacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactac 180
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30073 CAAAACGATTCGTATTTATATGATGACATCGAAATTCCTGCTAGCTTAACTGGATACTAC 30132

Qy    181 gagttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgt 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30133 GAGTTCAGGCAGTTTGGGGACGGTTCGATTACGCCGATTGAAAAGCATTTAAGGGCCTGT 30192

Qy    241 gtttgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttggccaac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30193 GTTTGCAGCGTGCGGCCCTGTATTCGAATCTGCTGCCCAGCCAAGAACTTTTGGCCAAC 30252

Qy    301 ggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30253 GGAAAATGCGATGATGGTCTCAAAGAGGAGCTCGCCCGGTTCAAACCCTATATATACTTC 30312

Qy    361 acatacatggacctacaggcacgagtaccactcaccgatatggctattatcagagatgag 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30313 ACATACATGGACCTACAGGCACGAGTACCACTCACCGATATGGCTATTATCAGAGATGAG 30372

Qy    421 ttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggtt 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30373 TTCTTTGATTGTGATGAAATGATTTACATCAGCGACTTTAACTATTTCTTAGAAGAGGTT 30432

Qy    481 agcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaatttttgg 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30433 AGCATTCAAATCTTCAATAAATGTGGACTTATAGTTTGGTTCCAGGATGGAAAATTTTGG 30492

Qy    541 gttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaactttgat 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30493 GTTACCGTTGACCTCTTCATGGAAAAACAGGACTACTGCTTGTACCGGCACAACTTTGAT 30552

Qy    601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30553 TCGGATTTTCCAAAGTCCATGTGGATAATACGACACCGCTGCACAAGCCACATATCTCCT 30612

Qy    661 ggatccttagaga 673
          |||||||||||||
Db  30613 GGATCCTTAGAGA 30625


RESULT   2
AC020432
LOCUS       AC020432    83545 bp    DNA             HTG       03-JAN-2000
DEFINITION  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
            pieces.
ACCESSION   AC020432
VERSION     AC020432.1  GI:6664465
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 83545)
  AUTHORS   Adams,M. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     This sequence was identified as CDM:10213149 by the submitter.
            For more information on this record e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .83545
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
BASE COUNT    24666 a  18278 c  17979 g  22622 t
ORIGIN


  Query Match             43.8%;  Score 673;  DB 65;  Length 83545;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  673;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 atgaggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33705 ATGAGGCTTCCTTGGGTGATTTTCTGTACAGTTCTGCTTTTAATTTTTACGAACAACTCA 33764

Qy     61 aatgcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaagg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33765 AATGCAGATATTCCCGGGTGCAACTACTACGACACGGTTGATATCTCATACATTGAAAGG 33824

Qy    121 caaaacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactac 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33825 CAAAACGATTCGTATTTATATGATGACATCGAAATTCCTGCTAGCTTAACTGGATACTAC 33884

Qy    181 gagttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgt 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33885 GAGTTCAGGCAGTTTGGGGACGGTTCGATTACGCCGATTGAAAAGCATTTAAGGGCCTGT 33944

Qy    241 gtttgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttggccaac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33945 GTTTGCAGCGTGCGGCCCTGTATTCGAATCTGCTGCCCAGCCAAGAACTTTTGGCCAAC 34004

Qy    301 ggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  34005 GGAAAATGCGATGATGGTCTCAAAGAGGAGCTCGCCCGGTTCAAACCCTATATATACTTC 34064

Qy    361 acatacatggacctacaggcacgagtaccactcaccgatatggctattatcagagatgag 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  34065 ACATACATGGACCTACAGGCACGAGTACCACTCACCGATATGGCTATTATCAGAGATGAG 34124

Qy    421 ttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggtt 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  34125 TTCTTTGATTGTGATGAAATGATTTACATCAGCGACTTTAACTATTTCTTAGAAGAGGTT 34184

Qy    481 agcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaatttttgg 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  34185 AGCATTCAAATCTTCAATAAATGTGGACTTATAGTTTGGTTCCAGGATGGAAAATTTTGG 34244

Qy    541 gttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaactttgat 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  34245 GTTACCGTTGACCTCTTCATGGAAAAACAGGACTACTGCTTGTACCGGCACAACTTTGAT 34304

Qy    601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  34305 TCGGATTTTCCAAAGTCCATGTGGATAATACGACACCGCTGCACAAGCCACATATCTCCT 34364

Qy    661 ggatccttagaga 673
          |||||||||||||
Db  34365 GGATCCTTAGAGA 34377


RESULT   3
AE003556/c
LOCUS       AE003556   281146 bp    DNA             INV       06-OCT-2000
DEFINITION  Drosophila melanogaster genomic scaffold 142000013386050 section 12
            of 54, complete sequence.
ACCESSION   AE003556 AE002602
VERSION     AE003556.2  GI:10728058
KEYWORDS    HTG.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 281146)
  AUTHORS   Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
            Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
            George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
            Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
            Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
            Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
            Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
```

```
            Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
            Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
            Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
            Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
            Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
            Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
            Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
            Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
            Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S.,
            Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S.,
            Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
            Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
            Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
            Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
            Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
            Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
            Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
            McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
            Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
            Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
            Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
            Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
            Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
            Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
            Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,
            Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
            Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,
            Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
            Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
            Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
            Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
  TITLE     The genome sequence of Drosophila melanogaster
  JOURNAL   Science 287 (5461), 2185-2195 (2000)
  MEDLINE   20196006
REFERENCE   2  (bases 1 to 281146)
  AUTHORS   Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     On Oct 9, 2000 this sequence version replaced gi:7295117.
FEATURES             Location/Qualifiers
     source          1. .281146
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="3L"
     mRNA            complement(join(2028. .3137,3574. .3871,5578. .5744,
                     6159. .6388))
                     /gene="bip1"
                     /note="Nucleotide sequence of the Celera sequence differs
                     from the published sequence for this transcript."
                     /product="CT23173"
                     /db_xref="FLYBASE:FBan0007574"
                     /db_xref="FLYBASE:FBgn0026263"
     gene            complement(<2028. .>6388)
                     /gene="bip1"
                     /note="CG7574"
                     /map="66B4-66B5"
                     /db_xref="FLYBASE:FBan0007574"
                     /db_xref="FLYBASE:FBgn0026263"
     CDS             complement(join(2343. .3137,3574. .3871,5578. .5744))
                     /gene="bip1"
                     /note="bip1 gene product; Nucleotide sequence of the
                     Celera sequence differs from the published sequence for
                     this transcript"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0007574"
                     /db_xref="FLYBASE:FBgn0026263"
                     /protein_id="AAF50496.1"
                     /db_xref="GI:7295171"
                     /translation="MTNTITTLRICQAVSLNPSANGPEPVLPPSSRICLMGCQDVKPN
                     DLRRFPVHDTDRCNHWLRHFGVSYKLVDQLGGLQKLRICSRHFSQRQEMAKTRIKMTP
                     LRKSSSSKGIVLPINAAGGSVIAGALAGGGGAGTGSGAVKSAVASSTTIELVESPNRH
                     QAKKPNTGIMPMPQRSHSISSSASSDVLSINSDYEASLPLSALEDSQNGLNRLENGQA
                     KKRKLYVITEKSDETKGQTLETNGQPKKRKMFVVVEQDKDVAGKKLMIMADKPQTNLT
                     QHLEKLLPELLNRIQDKEQKQIKPAPVVTKAKPVVNNPRPHITLPPKKNPTNFAKITS
                     TPSSVPSSIQMPLPIQNTKILKSENSELTMNSDEIETNINLSPDFRFLMKILPKLEQL
                     PEPHKQNVKRSIQIFVEKSYSIYGNKD"
     mRNA            <19498. .>20010
                     /gene="CG13681"
                     /product="CT33117"
                     /db_xref="FLYBASE:FBan0013681"
                     /db_xref="FLYBASE:FBgn0035832"
                     /evidence=not_experimental
     gene            <19498. .>20010
                     /gene="CG13681"
                     /map="66B5-66B6"
                     /db_xref="FLYBASE:FBan0013681"
                     /db_xref="FLYBASE:FBgn0035832"
                     /evidence=not_experimental
     CDS             19498. .20010
                     /gene="CG13681"
                     /note="CG13681 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0013681"
                     /db_xref="FLYBASE:FBgn0035832"
                     /evidence=not_experimental
                     /protein_id="AAF50495.1"
                     /db_xref="GI:7295170"
                     /translation="MNSPQLAPTTAQSHHAESADKTAPPGNKYAPGRRLAARQTLRLA
                     IPKAQPGDPLNHPTSVGLFRPSPSQLPPAPILKRVPAVSGGCSGIPKSASPLESHHLP
                     NVTPTNAHTRLGYGMQGAGGGSSCSSASSASNFLGVRKPSLTLNSSDIRNNNDLFAFP
                     KDMSPVMSGK"
     mRNA            complement(join(35698. .35900,35969. .36136,36196. .36712,
                     36772. .36962,37025. .37133,37194. .37325,37382. .38665,
                     38732. .>39472))
                     /gene="CG7565"
                     /product="CT23145"
                     /db_xref="FLYBASE:FBan0007565"
                     /db_xref="FLYBASE:FBgn0035833"
     gene            complement(<35698. .>39472)
                     /gene="CG7565"
                     /map="66B6-66B7"
                     /db_xref="FLYBASE:FBan0007565"
                     /db_xref="FLYBASE:FBgn0035833"
     CDS             complement(join(35698. .35900,35969. .36136,36196. .36712,
                     36772. .36962,37025. .37133,37194. .37325,37382. .38665,
                     38732. .39337))
                     /gene="CG7565"
                     /note="CG7565 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0007565"
                     /db_xref="FLYBASE:FBgn0035833"
                     /protein_id="AAF50494.1"
                     /db_xref="GI:7295169"
                     /translation="MVQVGKRICNLLLLATAMSSAYADVTTQNALLVGSKKHKETSPD
                     NSVGGSISPNLVCHKMLRHVFENATPRDEQQAGVFEEYKPPPDAVEPLEEEAYLWNCL
                     QACCEKPRNGSSACNVVLVFKAKCYHIRCQSNEACLPKLRVRMPNEKVQMVLVNPLGD
                     ATWPQLLKAEAAKQNAEILPYDEAALNFWKQPRRLSYLARNQETPVYEDEDFPLADKR
                     MNQMIFQPDENDVLANEELGYYDSNAKFTTCDMETPCPPPQQCVPLQPNAVRGVCTCP
                     EGFVWNKQRKCVMAAVPYSSYLTSNEAGQQEAAASENSPEVSTPPLKAEQNKDIVVSV
                     MSKEVRLPEQEVTLAAFTVPDEQTSDTKYKYLWTLISQPKGPMNGTISDQSKSKVKLS
                     NLSEGLYTFKVTVTGDNGTFGEATANVTVLPENRINQPPQVIISPREQIIRQPTTNAI
                     LDGSTSTDDDKITNWHWEVISGPIGYQPVLPEVNTLQLDLTSPGNYTFKLTVTDSNNV
                     TNSTTATIAVLKETDYAPVANAGDAVILYLPNNNVTLNGTASSDDHEIVAWEWTKDAS
                     DEAKAVDMQNTRTPYVQLSNLEEGMYTFVLKVTDGSGQSSTAKVHVFVKPPTNSPPVA
                     EAGSNTTTSLPINWVLLNGSDSKDDIGIKSYLWKQLSGPNNAVILKSNSSIANATSLT
                     LGLYEFELTVADENNNTATDTTWVKIVQERNAAPIANAGGDHTVTLPATAIYFNGSKS
                     WDDLAVVKYLWTRDEHSLAAGVIVADTDKEPVMILTNLVQGRYVFTLTVSDDQGLTSS
                     DTVSVNVRRDPKLLNLVQMTLPMGISVLVQSELDSVVQKLQLLLGDENKIQVRELKYD
                     LHTDATVLVFYVNDGQGKALDGLQVERQLRTQLQKDASILGAFAVDIRTSVCQSDCSG
                     HGSCNPITRACICEAFWMPSAGYFFNNQEANCDWSILYVFVGVIVGCLLLSGVFWGIA
                     CACRQSKKPRLRQKVQKYSLIGNKDEEAANYSRNTSLTESETDSDVLFETRTKSNGLG
                     KHKSHNSHSHGHGSSGGSGSSGRDGHKYGTQKLGRKIKA"
     mRNA            join(<40759. .41037,41176. .>41187)
                     /gene="CG13680"
                     /product="CT33116"
```

```
                     /db_xref="FLYBASE:FBan0013680"
                     /db_xref="FLYBASE:FBgn0040834"
                     /evidence=not_experimental
     gene            <40759. .>41187
                     /gene="CG13680"
                     /map="66B6-66B7"
                     /db_xref="FLYBASE:FBan0013680"
                     /db_xref="FLYBASE:FBgn0040834"
                     /evidence=not_experimental
     CDS             join(40759. .41037,41176. .41187)
                     /gene="CG13680"
                     /note="CG13680 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0013680"
                     /db_xref="FLYBASE:FBgn0040834"
                     /evidence=not_experimental
                     /protein_id="AAF50493.1"
                     /db_xref="GI:7295168"
                     /translation="MTDFSDSSPSGVLSPPTSLIENFHKKNSSPQPLGSSSAISTLTP
                     TTLNGSVVGGGVGGAGHHSIENGRQINTSSSSCSSLKETSHQSTSTSQQNTD"
     mRNA            join(44247. .44576,44671. .44872,46046. .46425,46846. .47236,
                     47295. .>47339)
                     /gene="CG7905"
                     /product="CT23884"
                     /db_xref="FLYBASE:FBan0007905"
                     /db_xref="FLYBASE:FBgn0035834"
     gene            <44247. .>47339
                     /gene="CG7905"
                     /map="66B6-66B7"
                     /db_xref="FLYBASE:FBan0007905"
                     /db_xref="FLYBASE:FBgn0035834"
     CDS             join(46169. .46425,46846. .47236,47295. .47339)
                     /gene="CG7905"
                     /note="CG7905 gene product"
                     /codon_start=1

  Query Match             43.8%;  Score 673;  DB 4;  Length 281146;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  673;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 atgaggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 128217 ATGAGGCTTCCTTGGGTGATTTTCTGTACAGTTCTGCTTTTAATTTTTACGAACAACTCA 128158

Qy      61 aatgcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaagg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 128157 AATGCAGATATTCCCGGGTGCAACTACTACGACACGGTTGATATCTCATACATTGAAAGG 128098

Qy     121 caaaacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactac 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 128097 CAAAACGATTCGTATTTATATGATGACATCGAAATTCCTGCTAGCTTAACTGGATACTAC 128038

Qy     181 gagttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgt 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 128037 GAGTTCAGGCAGTTTGGGGACGGTTCGATTACGCCGATTGAAAAGCATTTAAGGGCCTGT 127978

Qy     241 gtttgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttggccaac 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127977 GTTTGCAGCGTGCGGCCCTGTATTCGAATCTGCTGCCCAGCCAAGAACTTTTGGCCAAC 127918

Qy     301 ggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttc 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127917 GGAAAATGCGATGATGGTCTCAAAGAGGAGCTCGCCCGGTTCAAACCCTATATATACTTC 127858

Qy     361 acatacatggacctacaggcacgagtaccactcaccgatatggctattatcagagatgag 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127857 ACATACATGGACCTACAGGCACGAGTACCACTCACCGATATGGCTATTATCAGAGATGAG 127798

Qy     421 ttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggtt 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127797 TTCTTTGATTGTGATGAAATGATTTACATCAGCGACTTTAACTATTTCTTAGAAGAGGTT 127738

Qy     481 agcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaatttttgg 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127737 AGCATTCAAATCTTCAATAAATGTGGACTTATAGTTTGGTTCCAGGATGGAAAATTTTGG 127678

Qy     541 gttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaacttttgat 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127677 GTTACCGTTGACCTCTTCATGGAAAAACAGGACTACTGCTTGTACCGGCACAACTTTGAT 127618

Qy     601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127617 TCGGATTTTCCAAAGTCCATGTGGATAATACGACACCGCTGCACAAGCCACATATCTCCT 127558

Qy     661 ggatccttagaga 673
           |||||||||||||
Db 127557 GGATCCTTAGAGA 127545


RESULT   4
AC010037/c
LOCUS       AC010037   164659 bp    DNA             HTG       29-MAY-2000
DEFINITION  Drosophila melanogaster clone RPCI98-6E4, WORKING DRAFT SEQUENCE,
            46 unordered pieces.
ACCESSION   AC010037
VERSION     AC010037.5  GI:8101304
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 164659)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
            Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
            Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
            Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
            Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
            Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
            Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M.,
            Watlington,S., Weinstock,G., Weinstock,I.R., Williamson,A.,
            Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 164659)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1999) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 29, 2000 this sequence version replaced gi:6996724.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: DRCO
                Center clone name: RPCI98-6E4
            -------------- Summary Statistics
                Sequencing vector: M13; L08821
                Chemistry: Dye-primer Bodipy: 29% of reads
                Chemistry: Dye-terminator Big Dye: 25% of reads
                Assembly program: Phrap; version 0.980611
                Consensus quality: 112955 bases at least Q40
                Consensus quality: 133293 bases at least Q30
```

```
    Consensus quality: 144574 bases at least Q20
    Estimated insert size: 160113; sum-of-contigs estimation
    Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 46 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*       1     6150: contig of 6150 bp in length
*    6151     6250: gap of unknown length
*    6251    13075: contig of 6825 bp in length
*   13076    13175: gap of unknown length
*   13176    18980: contig of 5805 bp in length
*   18981    19080: gap of unknown length
*   19081    24693: contig of 5613 bp in length
*   24694    24793: gap of unknown length
*   24794    33330: contig of 8537 bp in length
*   33331    33430: gap of unknown length
*   33431    40699: contig of 7269 bp in length
*   40700    40799: gap of unknown length
*   40800    46106: contig of 5307 bp in length
*   46107    46206: gap of unknown length
*   46207    50498: contig of 4292 bp in length
*   50499    50598: gap of unknown length
*   50599    55756: contig of 5158 bp in length
*   55757    55856: gap of unknown length
*   55857    59777: contig of 3921 bp in length
*   59778    59877: gap of unknown length
*   59878    66349: contig of 6472 bp in length
*   66350    66449: gap of unknown length
*   66450    71183: contig of 4734 bp in length
*   71184    71283: gap of unknown length
*   71284    76459: contig of 5176 bp in length
*   76460    76559: gap of unknown length
*   76560    81902: contig of 5343 bp in length
*   81903    82002: gap of unknown length
*   82003    86768: contig of 4766 bp in length
*   86769    86868: gap of unknown length
*   86869    92357: contig of 5489 bp in length
*   92358    92457: gap of unknown length
*   92458    97795: contig of 5338 bp in length
*   97796    97895: gap of unknown length
*   97896   101452: contig of 3557 bp in length
*  101453   101552: gap of unknown length
*  101553   105177: contig of 3625 bp in length
*  105178   105277: gap of unknown length
*  105278   108706: contig of 3429 bp in length
*  108707   108806: gap of unknown length
*  108807   111390: contig of 2584 bp in length
*  111391   111490: gap of unknown length
*  111491   113983: contig of 2493 bp in length
*  113984   114083: gap of unknown length
*  114084   116768: contig of 2685 bp in length
*  116769   116868: gap of unknown length
*  116869   119023: contig of 2155 bp in length
*  119024   119123: gap of unknown length
*  119124   121906: contig of 2783 bp in length
*  121907   122006: gap of unknown length
*  122007   124731: contig of 2725 bp in length
*  124732   124831: gap of unknown length
*  124832   126964: contig of 2133 bp in length
*  126965   127064: gap of unknown length
*  127065   129583: contig of 2519 bp in length
*  129584   129683: gap of unknown length
*  129684   132383: contig of 2700 bp in length
*  132384   132483: gap of unknown length
*  132484   134455: contig of 1972 bp in length
*  134456   134555: gap of unknown length
*  134556   136568: contig of 2013 bp in length
*  136569   136668: gap of unknown length
*  136669   138348: contig of 1680 bp in length
*  138349   138448: gap of unknown length
*  138449   141263: contig of 2815 bp in length
*  141264   141363: gap of unknown length
*  141364   143990: contig of 2627 bp in length
*  143991   144090: gap of unknown length
*  144091   146194: contig of 2104 bp in length
*  146195   146294: gap of unknown length
*  146295   148427: contig of 2133 bp in length
*  148428   148527: gap of unknown length
*  148528   150406: contig of 1879 bp in length
*  150407   150506: gap of unknown length
*  150507   152893: contig of 2387 bp in length
*  152894   152993: gap of unknown length
*  152994   154351: contig of 1358 bp in length
*  154352   154451: gap of unknown length
*  154452   156305: contig of 1854 bp in length
*  156306   156405: gap of unknown length
*  156406   157843: contig of 1438 bp in length
*  157844   157943: gap of unknown length
*  157944   159199: contig of 1256 bp in length
*  159200   159299: gap of unknown length
*  159300   160524: contig of 1225 bp in length
*  160525   160624: gap of unknown length
*  160625   161884: contig of 1260 bp in length
*  161885   161984: gap of unknown length
*  161985   163299: contig of 1315 bp in length
*  163300   163399: gap of unknown length
*  163400   164659: contig of 1260 bp in length.
FEATURES             Location/Qualifiers
     source          1. .164659
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RPCI98-6E4"
     misc_feature    complement(16179. .16278)
                     /note="assembly_fragment:contig14;
                     vector:Vector_pBACe3.6;
                     vector_side:right"
     misc_feature    complement(53461. .53560)
                     /note="assembly_fragment:contig29;
                     vector:Vector_pBACe3.6;
                     vector_side:left"
BASE COUNT    44679 a  34729 c  35141 g  45537 t   4573 others
ORIGIN


  Query Match             28.4%;  Score 436;  DB 61;  Length 164659;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-252;
  Matches  436;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1101 agattggcacccatccgtctggatctatataagtggaccgtcgctggccctgagcacttt 1160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   59662 AGATTGGCACCCATCCGTCTGGATCTATATAAGTGGACCGTCGCTGGCCCTGAGCACTTT 59603

Qy    1161 caatgtcgccatgttcgccctgacagccatttacattaggaaagtgaaggggggtataaa 1220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   59602 CAATGTCGCCATGTTCGCCCTGACAGCCATTTACATTAGGAAAGTGAAGGGGGGTATAAA 59543

Qy    1221 taagttcacaaatgaggaggagggaaggattaactgcataaactttgacagccagacgta 1280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   59542 TAAGTTCACAAATGAGGAGGAGGGAAGGATTAACTGCATAAACTTTGACAGCCAGACGTA 59483

Qy    1281 agatatgataaaaccttggaccttggtaaactagttatccttgccgtattatccacatag 1340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   59482 AGATATGATAAAACCTTGGACCTTGGTAAACTAGTTATCCTTGCCGTATTATCCACATAG 59423

Qy    1341 ttacctacagttcctgcggctctccatcgtgatgggccttacttggatattcaatgtcat 1400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   59422 TTACCTACAGTTCCTGCGGCTCTCCATCGTGATGGGCCTTACTTGGATATTCAATGTCAT 59363

Qy    1401 tccgtattctgcacggctccacattttctgggaatgggtcgggataatatccgagtattt 1460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  59362 TCCGTATTCTGCACGGCTCCACATTTTCTGGGAATGGGTCGGGATAATATCCGAGTATTT 59303

Qy   1461 tcacagcgcgtttggaattgttttgttcgttctgctcgtcctgaagcgcagcacatggac 1520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  59302 TCACAGCGCGTTTGGAATTGTTTTGTTCGTTCTGCTCGTCCTGAAGCGCAGCACATGGAC 59243

Qy   1521 tctaatgatggattct 1536
          ||||||||||||||||
Db  59242 TCTAATGATGGATTCT 59227


RESULT   5
AF053725
LOCUS       AF053725     4616 bp    DNA             INV       31-MAR-1998
DEFINITION  Drosophila melanogaster myristoyl-CoA: protein N-myristoyl
            transferase (Nmt) gene, complete cds.
ACCESSION   AF053725
VERSION     AF053725.1  GI:2996201
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 4616)
  AUTHORS   Ntwasa,M., Egerton,M. and Gay,N.J.
  TITLE     Sequence and expression of Drosophila myristoyl-CoA: protein
            N-myristoyl transferase: evidence for proteolytic processing and
            membrane localisation
  JOURNAL   J. Cell. Sci. 110 (Pt 2), 149-156 (1997)
  MEDLINE   97196956
REFERENCE   2  (bases 1 to 4616)
  AUTHORS   Ntwasa,M., Egerton,M. and Gay,N.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-1998) Anatomy and Cell Biology, University of
            Cape Town, Anzio Road, Observatory, Cape Town 7925, South Africa
FEATURES             Location/Qualifiers
     source          1. .4616
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="3"
                     /map="3L; 66C"
     gene            <1729. .>2892
                     /gene="Nmt"
     CDS             1729. .2892
                     /gene="Nmt"
                     /note="dNMT"
                     /codon_start=1
                     /product="myristoyl-CoA: protein N-myristoyl transferase"
                     /protein_id="AAC08578.1"
                     /db_xref="GI:2996202"
                     /translation="MASTRQMAKKFAFWSTQPVTKLDEQVTTNECIEPNKEISEIRAL
                     PYTLPGFKWVTLDLNDANDLKELYTLLNENYVEDDDAMFRFDYQPEFLKWSLQPPGWK
                     RDWHVGVRVEKSGKLVGFISAIPSKLKSYDKVLKVVDINFLCVHKKLRSKRVAPVLIR
                     EISRRVNLTGIFQAAYTAGVVLPTPVATCRYWHRSLNPKKLVDVRFSHLARNMTMQRT
                     MKLYKLPDQPKTKGYRRITAKDMDKAHKLLEDYLKRFQLSPVFSKEEFRHWFTPKEGI
                     IDCFVVADEKGNITDLTSYYCLPSSVMHHPVHKTVRAAYSFYNVSTKTPWLDLMGDAL
                     ISARNVQMDVYNALDLMENKKYFAPLKFGAGDGNLQYYLYNWRCPSMQPERLL"
BASE COUNT     1297 a   1064 c   1048 g   1207 t
ORIGIN


  Query Match             12.4%;  Score 191;  DB 5;  Length 4616;
  Best Local Similarity   99.6%;  Pred. No. 1.2e-103;
  Matches  241;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   1135 ggaccgtcgctggccctgagcactttcaatgtcgccatgttcgccctgacagccatttac 1194
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    256 GGACCGTCGCTGGCCCTGAGCACTTTCAATGTCGCCATGTTCGCCCTGACAGCCATTTAC 315

Qy   1195 attaggaaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggattaac 1254
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    316 ATTAGGAAAGTGAAGGGGGGTATAAATAAGTTCACAAATGAGGAGGAGGGAAGGATTAAC 375

Qy   1255 tgcataaactttgacagccagacgtaagatatgataaaaccttggaccttggtaaactag 1314
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    376 TGCATCAACTTTGACAGCCAGACGTAAGATATGATAAAACCTTGGACCTTGGTAAACTAG 435

Qy   1315 ttatccttgccgtattatccacatagttacctacagttcctgcggctctccatcgtgatg 1374
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    436 TTATCCTTGCCGTATTATCCACATAGTTACCTACAGTTCCTGCGGCTCTCCATCGTGATG 495

Qy   1375 gg 1376
          ||
Db    496 GG 497


RESULT   6
AX034794/c
LOCUS       AX034794     400 bp    DNA             PAT       15-NOV-2000
DEFINITION  Sequence 2 from Patent WO0051424.
ACCESSION   AX034794
VERSION     AX034794.1  GI:11190750
KEYWORDS    .
SOURCE      Ovis sp.
  ORGANISM  Ovis sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 400)
  AUTHORS   Kind,A.J., Schnieke,A.E., Colman,A., Dai,Y. and Ayares,D.L.
  TITLE     Genetic modification of somatic cells and uses thereof
  JOURNAL   Patent: WO 0051424-A 2 08-SEP-2000;
            KIND ALEXANDER JARVIS (GB) ; PPL THERAPEUTICS SCOTLAND LTD (GB) ;
            SCHNIEKE ANGELIKA ELISABETH (GB) ; COLMAN ALAN (GB) ; DAI YIFAN
            (US) ; AYARES DAVID LEE (US)
FEATURES             Location/Qualifiers
     source          1. .400
                     /organism="Ovis sp."
                     /db_xref="taxon:9939"
BASE COUNT       90 a    133 c     83 g     94 t
ORIGIN


  Query Match             1.4%;  Score 21;  DB 9;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    603 ggattttccaaagtccatgtg 623
          |||||||||||||||||||||
Db    171 GGATTTTCCAAAGTCCATGTG 151


RESULT   7
AC067792
LOCUS       AC067792   164232 bp    DNA             HTG       28-MAY-2000
DEFINITION  Homo sapiens chromosome 4 clone RP11-612A5 map 4, WORKING DRAFT
            SEQUENCE, 29 unordered pieces.
ACCESSION   AC067792
VERSION     AC067792.2  GI:8099932
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 164232)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 4, clone RP11-612A5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 164232)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
```

```
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 28, 2000 this sequence version replaced gi:7651842.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L9862
                Center clone name: 612_A_5
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 152206 bases at least Q40
                Consensus quality: 157744 bases at least Q30
                Consensus quality: 159901 bases at least Q20
                Insert size: 170000; agarose-fp
                Insert size: 161432; sum-of-contigs
                Quality coverage: 4.1 in Q20 bases; agarose-fp
                Quality coverage: 4.3 in Q20 bases; sum-of-contigs
            --------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 29 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1132: contig of 1132 bp in length
            *     1133 1232: gap of      100 bp
            *     1233     2360: contig of 1128 bp in length
            *     2361 2460: gap of      100 bp
            *     2461     3776: contig of 1316 bp in length
            *     3777 3876: gap of      100 bp
            *     3877     5502: contig of 1626 bp in length
            *     5503 5602: gap of      100 bp
            *     5603     7002: contig of 1400 bp in length
            *     7003 7102: gap of      100 bp
            *     7103     9509: contig of 2407 bp in length
            *     9510 9609: gap of      100 bp
            *     9610    11542: contig of 1933 bp in length
            *    11543 11642: gap of      100 bp
            *    11643    12988: contig of 1346 bp in length
            *    12989 13088: gap of      100 bp
            *    13089    13883: contig of 795 bp in length
            *    13884 13983: gap of      100 bp
            *    13984    16205: contig of 2222 bp in length
            *    16206 16305: gap of      100 bp
            *    16306    19854: contig of 3549 bp in length
            *    19855 19954: gap of      100 bp
            *    19955    22823: contig of 2869 bp in length
            *    22824 22923: gap of      100 bp
            *    22924    25386: contig of 2463 bp in length
            *    25387 25486: gap of      100 bp
            *    25487    28778: contig of 3292 bp in length
            *    28779 28878: gap of      100 bp
            *    28879    31839: contig of 2961 bp in length
            *    31840 31939: gap of      100 bp
            *    31940    36890: contig of 4951 bp in length
            *    36891 36990: gap of      100 bp
            *    36991    39944: contig of 2954 bp in length
            *    39945 40044: gap of      100 bp
            *    40045    43979: contig of 3935 bp in length
            *    43980 44079: gap of      100 bp
            *    44080    48513: contig of 4434 bp in length
            *    48514 48613: gap of      100 bp
            *    48614    54700: contig of 6087 bp in length
            *    54701 54800: gap of      100 bp
            *    54801    60314: contig of 5514 bp in length
            *    60315 60414: gap of      100 bp
            *    60415    67391: contig of 6977 bp in length
            *    67392 67491: gap of      100 bp
            *    67492    75618: contig of 8127 bp in length
            *    75619 75718: gap of      100 bp
            *    75719    85397: contig of 9679 bp in length
            *    85398 85497: gap of      100 bp
            *    85498    97404: contig of 11907 bp in length
            *    97405 97504: gap of      100 bp
            *    97505   112118: contig of 14614 bp in length
            *   112119 112218: gap of      100 bp
            *   112219   129378: contig of 17160 bp in length
            *   129379 129478: gap of      100 bp
            *   129479   145549: contig of 16071 bp in length
            *   145550 145649: gap of      100 bp
            *   145650   164232: contig of 18583 bp in length.
FEATURES             Location/Qualifiers
     source          1. .164232
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /map="4"
                     /clone="RP11-612A5"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1132
                     /note="assembly_fragment"
     misc_feature    1233. .2360
                     /note="assembly_fragment"
     misc_feature    2461. .3776
                     /note="assembly_fragment"
     misc_feature    3877. .5502
                     /note="assembly_fragment"
     misc_feature    5603. .7002
                     /note="assembly_fragment"
     misc_feature    7103. .9509
                     /note="assembly_fragment"
     misc_feature    9610. .11542
                     /note="assembly_fragment"
     misc_feature    11643. .12988
                     /note="assembly_fragment"
     misc_feature    13089. .13883
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:left"
     misc_feature    13984. .16205
                     /note="assembly_fragment"
     misc_feature    16306. .19854
                     /note="assembly_fragment"
     misc_feature    19955. .22823
                     /note="assembly_fragment"
     misc_feature    22924. .25386
                     /note="assembly_fragment"
     misc_feature    25487. .28778
                     /note="assembly_fragment"
     misc_feature    28879. .31839
```

```
                     /note="assembly_fragment"
     misc_feature    31940. .36890
                     /note="assembly_fragment"
     misc_feature    36991. .39944
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     misc_feature    40045. .43979
                     /note="assembly_fragment"
     misc_feature    44080. .48513
                     /note="assembly_fragment"
     misc_feature    48614. .54700
                     /note="assembly_fragment"
     misc_feature    54801. .60314
                     /note="assembly_fragment"
     misc_feature    60415. .67391
                     /note="assembly_fragment"
     misc_feature    67492. .75618
                     /note="assembly_fragment"
     misc_feature    75719. .85397
                     /note="assembly_fragment"
     misc_feature    85498. .97404
                     /note="assembly_fragment"
     misc_feature    97505. .112118
                     /note="assembly_fragment"
     misc_feature    112219. .129378
                     /note="assembly_fragment"
     misc_feature    129479. .145549
                     /note="assembly_fragment"
     misc_feature    145650. .164232
                     /note="assembly_fragment"
BASE COUNT    51426 a  30879 c  29567 g  49553 t   2807 others
ORIGIN

  Query Match             1.4%;  Score 21;  DB 73;  Length 164232;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    986 ggagcacagccgcaatcatga 1006
          ||||||||||||||||||||
Db  77049 GGAGCACAGCCGCAATCATGA 77069


RESULT   8
SCYLL021W/c
LOCUS       SCYLL021W     5354 bp    DNA             PLN       11-AUG-1997
DEFINITION  S.cerevisiae chromosome XII reading frame ORF YLL021w.
ACCESSION   Z73126 Y13138
VERSION     Z73126.1  GI:1360195
KEYWORDS    .
SOURCE      Saccharomyces cerevisiae.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 5354)
  AUTHORS   Goffeau,A. and Purnelle,B.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 5354)
  AUTHORS   MIPS.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAY-1996) Data collected by MIPS on behalf of the
            European yeast chromosome XII sequencing project. MIPS at the
            Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152
            Martinsried, FRG; E-mail: Mewes@mips.embnet.org
FEATURES             Location/Qualifiers
     source          1. .5354
                     /organism="Saccharomyces cerevisiae"
                     /db_xref="taxon:4932"
                     /chromosome="XII"
     gene            746. .5146
                     /gene="SPA2"
     CDS             746. .5146
                     /gene="SPA2"
                     /note="ORF YLL021w"
                     /codon_start=1
                     /protein_id="CAA97469.1"
                     /db_xref="GI:1360196"
                     /db_xref="SWISS-PROT:P23201"
                     /translation="MGTSSEVSLAHHRDIFHYYVSLKTFFEVTGENRDRSNSTRAQKA
                     RAKLLKLSSSQFYELSTDVSDELQRRIGEDANQPDYLLPKANFHMKRNQARQKLANLS
                     QTRFNDLLDDILFEIKRRGFDKDLDAPRPPLPQPMKQEVSKDSDDTARTSTNSSSVTQ
                     VAPNVSVQPSLVIPKMASIDWSSEEEEEEQVKEKPNEPEGKQTSMDEKKEAKPALNPI
                     VTDSDLPDSQVLARDITSMARTPTTTHKNYWDVNDSPIIKVDKDIDNEKGPEQLKSPE
                     VQRAENNNPNSEMEDKVKELTDLNSDLHLQIEDLNAKLASLTSEKEKEKKEEKEEKEK
                     EKNLKINYTIDESFQKELLSLNSQIGELSIENENLKQKISEFELHQKKNDNHNDLKIT
                     DGFISKYSSADGLIPAQYILNANNLIIQFTTRLSAVPIGDSTAISHQIGEELFQILSQ
                     LSNLISQLLLSADLLQYKDQVILLKASLSHAITSIRYFSVYGPVLIPKITVQAAVSEV
                     CFAMCNLIDSAKIKSDSNGESTTSNEGNRQVLEYSSPTATTPMTPTFPSTSGINMKKG
                     FINPRKPASFLNDVEEEESPVKPLKITQKAINSPIIRPSSSNGVPTTSRKPSGTGLFS
                     LMIDSSIAKNSSHKEDNDKYVSPIKAVTSASNSASSNISEIPKLTLPPQAKIGTVIPP
                     SENQVPNIKIENTEEDNKRSDITNEISVKPTSSIADKLKQFEQSSEKKSSPKENPIAK
                     EEMDSKPKLSNKFITSMNDVSTDDSSSDGNENDDADDDDDFTYMALKQTMKREGSKIE
                     KNNDSKLPANIVELDLHESPESVKIESPESIKEITSSEMSSEMPSSSLPKRLVEDVEP
                     SEMPEKGASVESVRKKNFQEPLGNVESPDMTQKVKSLGMTGKAVGPESDSRVESPGMT
                     GQIKSLNMAGKVVGPEADSRVESPGMKEQIKSLGMTGKITAQESIKSPEAARKLASSG
                     EVDKIESPRMVRESESLEAVGNTIPSNMTVKMESPNLKGNTVSEPQEIRRDIASSEPI
                     ENVDPPKVLKKIVFPKAVNRTGSPKSVEKTPSSATLKKSGLPEPNSQIVSPELAKNSP
                     LAPIKKNVELRETNKPHTETITSVEPTNKDANTSWRDADLNRTIKREEEDEDFDRVNH
                     NIQITGAYTKTGKIDYHKIPVDRKAKSEAEVHTSEEDIDESNNVNGKRADAQIHITER
                     KHAFVNPTENSQVKKTSHSPFLNSKPVQYENSESNGGINNHIKIKNTGETTAHDEKHY
                     SDDDDSSYQFVPMKHEEQEQEQNRSEEEESEDDDEEEEDSDFDVDTFDIENPDNTLSE
                     LLLYLEHQTMDVISTIQSLLTSIKKPQVTKGNLRGESNAINQVIGQMVDATSISMEQS
                     RNANLKKHGDWVVQSLRDCSRRMTILCQLTGDGILAKEKSDQDYADKNFKQRLAGIAF
                     DVAKCTKELVKTVEEASLKDEINYLNSKLK"
BASE COUNT     1991 a    976 c   1107 g   1280 t
ORIGIN

  Query Match             1.3%;  Score 20;  DB 15;  Length 5354;
  Best Local Similarity   100.0%;  Pred. No. 7.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    479 ttagcattcaaatcttcaat 498
          ||||||||||||||||||||
Db   1683 TTAGCATTCAAATCTTCAAT 1664


RESULT   9
SCSPA2G/c
LOCUS       SCSPA2G      6041 bp    DNA             PLN       19-MAR-1991
DEFINITION  S. cerevisiae SPA2 gene.
ACCESSION   X53731
VERSION     X53731.1  GI:4519
KEYWORDS    morphogenesis; SPA2 gene.
SOURCE      Saccharomyces cerevisiae.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 6041)
  AUTHORS   Snyder,M.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-1990) Snyder M.P., Yale University, Dept. of
            Biology, New Haven, CT 06511, USA
REFERENCE   2  (bases 1 to 6041)
  AUTHORS   Gehrung,S. and Snyder,M.
  TITLE     The SPA2 gene of Saccharomyces cerevisiae is important for
            pheromone-induced morphogenesis and efficient mating
  JOURNAL   J. Cell Biol. 111 (4), 1451-1464 (1990)
  MEDLINE   91009481
FEATURES             Location/Qualifiers
     source          1. .6041
                     /organism="Saccharomyces cerevisiae"
                     /strain="S288C"
                     /db_xref="taxon:4932"
                     /chromosome="XII, map postion 21"
                     /clone_lib="EMBL3A;Synder"
```

```
                     /clone="pSPA2-172"
     gene            735. .5135
                     /gene="SPA2"
     CDS             735. .5135
                     /gene="SPA2"
                     /codon_start=1
                     /product="SPA2 protein"
                     /protein_id="CAA37763.1"
                     /db_xref="GI:4520"
                     /db_xref="SWISS-PROT:P23201"
                     /translation="MGTSSEVSLAHHRDIFHYYVSLKTFFEVTGENRDRSNSTRAQKA
                     RAKLLKLSSSQFYELSTDVSDELQRRIGEDANQPDYLLPKANFHMKRNQARQKLANLS
                     QTRFNDLLDDILFEIKRRGFDKDLDAPRPPLPQPMKQEVSKDSDDTARTSTNSSSVTQ
                     VAPNVSVQPSLVIPKMASIDWSSEEEEEEQVKEKPNEPEGKQTSMDEKKEAKPALNPI
                     VTDSDLPDSQVLARDITSMARTPTTTHKNYWDVNDSPIIKVDKDIDNEKGPEQLKSPE
                     VQRAENNNPNSEMEDKVKELTDLNSDLHLQIEDLNAKLASLTSEKEKEKKEEKEEKEK
                     EKNLKINYTIDESFQKELLSLNSQIGELSIENENLKQKISEFELHQKKNDNHNDLKIT
                     DGFISKYSSADGLIPAQYILNANNLIIQFTTRLSAVPIGDSTAISHQIGEELFQILSQ
                     LSNLISQLLLSADLLQYKDQVILLKASLSHAITSIRYFSVYGPVLIPKITVQAAVSEV
                     CFAMCNLIDSAKIKSDSNGESTTSNEGNRQVLEYSSPTATTPMTPTFPSTSGINMKKG
                     FINPRKPASFLNDVEEEESPVKPLKITQKAINSPIIRPSSSNGVPTTSRKPSGTGLFS
                     LMIDSSIAKNSSHKEDNDKYVSPIKAVTSASNSASSNISEIPKLTLPPQAKIGTVIPP
                     SENQVPNIKIENTEEDNKRSDITNEISVKPTSSIADKLKQFEQSSEKKSSPKENPIAK
                     EEMDSKPKLSNKFITSMNDVSTDDSSSDGNENDDADDDDDFTYMALKQTMKREGSKIE
                     KNNDSKLPANIVELDLHESPESVKIESPESIKEITSSEMSSEMPSSSLPKRLVEDVEP
                     SEMPEKGASVESVRKKNFQEPLGNVESPDMTQKVKSLGMTGKAVGPESDSRVESPGMT
                     GQIKSLNMAGKVVGPEADSRVESPGMKEQIKSLGMTGKITAQESIKSPEAARKLASSG
                     EVDKIESPRMVRESESLEAVGNTIPSNMTVKMESPNLKGNTVSEPQEIRRDIASSEPI
                     ENVDPPKVLKKIVFPKAVNRTGSPKSVEKTPSSATLKKSGLPEPNSQIVSPELAKNSP
                     LAPIKKNVELRETNKPHTETITSVEPTNKDANTSWRDADLNRTIKREEEDEDFDRVNH
                     NIQITGAYTKTGKIDYHKIPVDRKAKSEAEVHTSEEDIDESNNVNGKRADAQIHITER
                     KHAFVNPTENSQVKKTSHSPFLNSKPVQYENSESNGGINNHIKIKNTGETTAHDEKHY
                     SDDDDSSYQFVPMKHEEQEQEQNRSEEEESEDDDEEEEDSDFDVDTFDIENPDNTLSE
                     LLLYLEHQTMDVISTIQSLLTSIKKPQVTKGNLRGESNAINQVIGQMVDATSISMEQS
                     RNANLKKHGDWVVQSLRDCSRRMTILCQLTGDGILAKEKSDQDYADKNFKQRLAGIAF
                     DVAKCTKELVKTVEEASLKDEINYLNSKLK"
BASE COUNT     2215 a    1090 c    1241 g    1495 t
ORIGIN


  Query Match             1.3%;  Score 20;  DB 14;  Length 6041;
  Best Local Similarity   100.0%;  Pred. No. 7.9;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    479 ttagcattcaaatcttcaat 498
          ||||||||||||||||||||
Db   1672 TTAGCATTCAAATCTTCAAT 1653


RESULT  10
SC32KBF/c
LOCUS       SC32KBF     32088 bp    DNA             PLN       13-FEB-1997
DEFINITION  S.cerevisiae 32kb DNA fragment of chromosome XII, left arm.
ACCESSION   X97560
VERSION     X97560.1  GI:1297003
KEYWORDS    ABC transporter; heat shock protein 104; hsp104 gene; KNS1 protein;
            SDC25 protein; seripauperins family; SPA2 protein; SSA2 protein.
SOURCE      Saccharomyces cerevisiae.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 32088)
  AUTHORS   Purnelle,B. and Goffeau,A.
  TITLE     The sequence of 32b on the left arm of yeast chromosome XII reveals
            six known genes, a new member of the seripauperins family and a new
            ABS transporter homologous to the human multidrug resistance
            protein
  JOURNAL   Yeast 13 (2), 183-188 (1997)
  MEDLINE   97197984
REFERENCE   2  (bases 1 to 32088)
  AUTHORS   Purnelle,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-1996) B. Purnelle, Unite de Biochimie
            Physiologique, Universite Catholique de Louvain, Place Croix du Sud
            2/20, 1348 Louvain-la-Neuve, BELGIUM
COMMENT     Related sequences: X91488, M67479, X12926, X53731, M85200, X06665,
            X03606 & M26647.
FEATURES             Location/Qualifiers
     source          1. .32088
                     /organism="Saccharomyces cerevisiae"
                     /strain="alpha S288C"
                     /db_xref="taxon:4932"
                     /chromosome="XII, left arm"
                     /clone_lib="cosmid Lawrist7 derivative"
                     /clone="2G12"
     gene            1. .484
                     /gene="hsp104"
     CDS             <1. .484
                     /gene="hsp104"
                     /note="L0948"
                     /codon_start=2
                     /product="heat shock protein 104"
                     /protein_id="CAA66164.1"
                     /db_xref="GI:1297004"
                     /db_xref="SWISS-PROT:P31539"
                     /translation="LVMGAVRQHFRPEFLNRISSIVIFNKLSRKAIHKIVDIRLKEIE
                     ERFEQNDKHYKLNLTQEAKDFLAKYGYSDDMGARPLNRLIQNEILNKLALRILKNEIK
                     DKETVNVVLKKGKSRDENVPEEAEECLEVLPNHEATIGADTLGDDDNEDSMEIDDDLD
                     "
     LTR             complement(1185. .1509)
                     /note="L0953"
     tRNA            join(1683. .1721,1753. .1782)
                     /gene="tRNA-Pro"
                     /note="L0690"
                     /product="transfer RNA-Pro"
     gene            join(1683. .1721,1753. .1782)
                     /gene="tRNA-Pro"
     CDS             3882. .4256
                     /note="hypothetical protein L0968
                     new member of seripauperins family"
                     /codon_start=1
                     /protein_id="CAA66166.1"
                     /db_xref="GI:1297006"
                     /db_xref="SPTREMBL:Q12370"
                     /translation="MVKLTSIAAGVAAIAAGVAAAPATTTLSPSDERVNLVELGVYVS
                     DIRAHLAEYYMFQAAHPTETYPVEIAEAVFNYGDFTTMLTGIPADQVTRVITGVPWYS
                     TRLRPAISSALSADGIYTAVPN"
     CDS             complement(4701. .6620)
                     /note="L0931
                     SSA2 protein"
                     /codon_start=1
                     /product="heat shock protein"
                     /protein_id="CAA66167.1"
                     /db_xref="GI:1297007"
                     /db_xref="SWISS-PROT:P10592"
                     /translation="MSKAVGIDLGTTYSCVAHFSNDRVDIIANDQGNRTTPSFVGFTD
                     TERLIGDAAKNQAAMNPANTVFDAKRLIGRNFNDPEVQGDMKHFPFKLIDVDGKPQIQ
                     VEFKGETKNFTPEQISSMVLGKMKETAESYLGAKVNDAVVTVPAYFNDSQRQATKDAG
                     TIAGLNVLRIINEPTAAAIAYGLDKKGKEEHVLIFDLGGGTFDVSLLSIEDGIFEVKA
                     TAGDTHLGGEDFDNRLVNHFIQEFKRKNKKDLSTNQRALRRLRTACERAKRTLSSSAQ
                     TSVEIDSLFEGIDFYTSITRARFEELCADLFRSTLDPVEKVLRDAKLDKSQVDEIVLV
                     GGSTRIPKVQKLVTDYFNGKEPNRSINPDEAVAYGAAVQAAILTGDESSKTQDLLLLD
                     VAPLSLGIETAGGVMTKLIPRNSTIPTKKSEVFSTYADNQPGVLIQVFEGERAKTKDN
                     NLLGKFELSGIPPAPRGVPQIEVTFDVDSNGILNVSAVEKGTGKSNKITITNDKGRLS
                     KEDIEKMVAEAEKFKEEDEKESQRIASKNQLESIAYSLKNTISEAGDKLEQADKDAVT
                     KKAEETIAWLDSNTTATKEEFDDQLKELQEVANPIMSKLYQAGGAPEGAAPGGFPGGA
                     PPAPEAEGPTVEEVD"
     CDS             complement(7132. .7971)
                     /note="hypothetical protein L1201"
                     /codon_start=1
                     /protein_id="CAA66168.1"
                     /db_xref="GI:1297008"
                     /db_xref="SPTREMBL:Q12164"
                     /translation="MSSRPANNQGPPNLPARDKSLVQRFMAVAKSLQFAWFTGHSVVL
                     ISSILYLLKMSEFYYRSAYLGVIESFGIIIYQQFFTRNEPLQTQDAAATKASIKSRVA
                     GLLKSEDVLYLVLANFWLFTPRFSFSLIPFFAFAVFHVLIYVEKVLLPKVFHLSSKDS
```

SKILSFIDKFVVQYNDLCMHWVGTAELLIFILVLFRAILCFQRSWIILVVYAIFIKLR
YENSKYMKAAFAQWRVRMDGIISHPSIPPFVKRAYNAIKMSLIRLSEYRLSGAPQVTK
KQN"

CDS complement(8179. .9336)
/note="hypothetical protein L1205"
/codon_start=1
/protein_id="CAA66169.1"
/db_xref="GI:1297009"
/db_xref="SPTREMBL:Q12373"
/translation="MKLRAEDVLANGTSRHKVQIDMERQVQIAKDLLAQKKFLEAAKR
CQQTLDSLPKDGLLPDPELFTIFAQAVYNMEVQNSGNLFGDALLAGDDGSGSESESEP
ESDVSNGEEGNENGQTEIPNSRMFQFDQEEEDLTGDVDSGDSEDSGEGSEEEEENVEK
EEERLALHELANFSPANEHDDEIEDVSQLRKSGPHIYFENDLYENALDLLAQALMLLG
RPTADGQSLTENSRLRIGDVYILMGDIEREAEMFSRAIHHYLKALGYYKTLKPAEQVT
EKVIQAEFLVCDALRWVDQVPAKDKLKRFKHAKALLEKHMTTRPKDSELQQARLAQIQ
DDIDEVQENQQHGSKRPLSQPTTSIGFPALEKPLGDFNDLSQLVKKKPRRH"

CDS 10082. .14482
/function="involved in cellular morphogenesis during
mating"
/note="SPA2 protein
L1209"
/codon_start=1
/protein_id="CAA66170.1"
/db_xref="GI:1297010"
/db_xref="SWISS-PROT:P23201"
/translation="MGTSSEVSLAHHRDIFHYYVSLKTFFEVTGENRDRSNSTRAQKA
RAKLLKLSSSQFYELSTDVSDELQRRIGEDANQPDYLLPKANFHMKRNQARQKLANLS
QTRFNDLLDDILFEIKRRGFDKDLDAPRPPLPQPMKQEVSKDSDDTARTSTNSSSVTQ
VAPNVSVQPSLVIPKMASIDWSSEEEEEEQVKEKPNEPEGKQTSMDEKKEAKPALNPI
VTDSDLPDSQVLARDITSMARTPTTTHKNYWDVNDSPIIKVDKDIDNEKGPEQLKSPE
VQRAENNNPNSEMEDKVKELTDLNSDLHLQIEDLNAKLASLTSEKEKEKKEEKEEKEK
EKNLKINYTIDESFQKELLSLNSQIGELSIENENLKQKISEFELHQKKNDNHNDLKIT
DGFISKYSSADGLIPAQYILNANNLIIQFTTRLSAVPIGDSTAISHQIGEELFQILSQ
LSNLISQLLLSADLLQYKDQVILLKASLSHAITSIRYFSVYGPVLIPKITVQAAVSEV
CFAMCNLIDSAKIKSDSNGESTTSNEGNRQVLEYSSPTATTPMTPTFPSTSGINMKKG
FINPRKPASFLNDVEEEESPVKPLKITQKAINSPIIRPSSSNGVPTTSRKPSGTGLFS
LMIDSSIAKNSSHKEDNDKYVSPIKAVTSASNSASSNISEIPKLTLPPQAKIGTVIPP
SENQVPNIKIENTEEDNKRSDITNEISVKPTSSIADKLKQFEQSSEKKSSPKENPIAK
EEMDSKPKLSNKFITSMNDVSTDDSSSDGNENDDADDDDDFTYMALKQTMKREGSKIE
KNNDSKLPANIVELDLHESPESVKIESPESIKEITSSEMSSEMPSSSLPKRLVEDVEP
SEMPEKGASVESVRKKNFQEPLGNVESPDMTQKVKSLGMTGKAVGPESDSRVESPGMT
GQIKSLNMAGKVVGPEADSRVESPGMKEQIKSLGMTGKITAQESIKSPEAARKLASSG
EVDKIESPRMVRESESLEAVGNTIPSNMTVKMESPNLKGNTVSEPQEIRRDIASSEPI
ENVDPPKVLKKIVFPKAVNRTGSPKSVEKTPSSATLKKSGLPEPNSQIVSPELAKNSP
LAPIKKNVELRETNKPHTETITSVEPTNKDANTSWRDADLNRTIKREEEDEDFDRVNH
NIQITGAYTKTGKIDYHKIPVDRKAKSEAEVHTSEEDIDESNNVNGKRADAQIHITER
KHAFVNPTENSQVKKTSHSPFLNSKPVQYENSESNGGINNHIKIKNTGETTAHDEKHY
SDDDDSSYQFVPMKHEEQEQEQNRSEEEESEDDDEEEEDSDFDVDTFDIENPDNTLSE
LLLYLEHQTMDVISTIQSLLTSIKKPQVTKGNLRGESNAINQVIGQMVDATSISMEQS
RNANLKKHGDWVVQSLRDCSRRMTILCQLTGDGILAKEKSDQDYADKNFKQRLAGIAF
DVAKCTKELVKTVEEASLKDEINYLNSKLK"

CDS complement(14821. .17034)
/note="KNS1 protein
L124"
/codon_start=1
/protein_id="CAA66171.1"
/db_xref="GI:1297011"
/db_xref="SPTREMBL:Q12399"
/translation="MSQNIQIGTRKRSRANMNNSTTTGPANNTSSNKTFLDNFEETRT
NKLLDEMFARQNSFLTDNLRNSLDLNQADNPLRPRQHQHQLFLDNENAIELDEEPRII
NTTINNSNNHNSSRVDEDADDDIIFIKEQPIQFSSPLILPSSSSINNNNNIVTSNNPG
CGTAATSNSTYITTPKKFKKQRTISLPQLPLSKLSYQSNYFNVPDQTNAIVPRVTQTE
NELLHLTGSCAKTLEGNKAVNLTIAHSTSPFSNPPAQIASLPQSNLKKQIGSSLRKFT
SNGSSESASSNKSNFKTDKDGHYVYQENDIFGSGGRFVVKDLLGQGTFGKVLKCIDNK
YEPNYVAVKVIRAVDRYREAAKTELRILQTILNNDPQGQFQCLLLRECFDYKNHICLV
TDLYGRSIYDFMCSNGIARFPGSHIQAIARQLIRSVCFLHDLGIIHTDLKPENILICD
ETHIAQKLPLKTVQSLSKRRREASKGKRKILKNPEIKIIDFGSAIFHYEYHPPVISTR
HYRAPEIVLGLGWSFPCDIWSIACVLVELVIGESLYPIHENLEHMAMMQRINGTPFPT
DIIDKMFYKSKHKLGNSPSDLNSTVIKHFDRKTLSLQWPEKNKRGDTITTEKSMKRVL
QSCDRLDIYISKVLKQDYGDSLSINWNLPPEKNWSLINSKLAWKRQTHSSSSSTTDEL
DKETFLFWYWFIDLLRKMFEFDPTKRITAKDALDHEWFNLGILDDGIATYNNTQG"

CDS complement(19037. .20710)
/note="DPS1/APS/APS1 protein
L1295"
/codon_start=1
/product="aspartyl-tRNA synthetase"
/protein_id="CAA66172.1"
/db_xref="GI:1297012"
/db_xref="SWISS-PROT:P04802"
/translation="MSQDENIVKAVEESAEPAQVILGEDGKPLSKKALKKLQKEQEKQ
RKKEERALQLEAEREAREKKAAAEDTAKDNYGKLPLIQSRDSDRTGQKRVKFVDLDEA
KDSDKEVLFRARVHNTRQQGATLAFLTLRQQASLIQGLVKANKEGTISKNMVKWAGSL
NLESIVLVRGIVKKVDEPIKSATVQNLEIHITKIYTISETPEALPILLEDASRSEAEA
EAAGLPVVNLDTRLDYRVIDLRTVTNQAIFRIQAGVCELFREYLATKKFTEVHTPKLL
GAPSEGGSSVFEVTYFKGKAYLAQSPQFNKQQLIVADFERVYEIGPVFRAENSNTHRH
MTEFTGLDMEMAFEEHYHEVLDTLSELFVFIFSELPKRFAHEIELVRKQYPVEEFKLP
KDGKMVRLTYKEGIEMLRAAGKEIGDFEDLSTENEKFLGKLVRDKYDTDFYILDKFPL
EIRPFYTMPDPANPKYSNSYDFFMRGEEILSGAQRIHDHALLQERMKAHGLSPEDPGL
KDYCDGFSYGCPPHAGGGIGLERVVMFYLDLKNIRRASLFPRDPKRLRP"

CDS 21369. .21680
/note="hypothetical protein L1305"
/codon_start=1
/protein_id="CAA66173.1"
/db_xref="GI:1297013"
/db_xref="SPTREMBL:Q12065"
/translation="MSCTASYAGMTTPVKDKEGHGIPCLQPIDVVECTYQYFTKSRNK
LSLRVGDLIYVLTKGSNGWWDGVLIRHSANNNNNSLILDRGWFPPSFYTVHSKRTTRG
A"

CDS 21982. .25128
/note="hypothetical protein L1309"
/codon_start=1
/protein_id="CAA66161.1"
/db_xref="GI:1297014"
/db_xref="SPTREMBL:Q12037"
/translation="MPITSSPDLFYLNDCDVVYWYDLTRLVCHYVNLTERDLLANERE

Query Match 1.3%; Score 20; DB 14; Length 32088;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 479 ttagcattcaaatcttcaat 498
||||||||||||||||||||
Db 11019 TTAGCATTCAAATCTTCAAT 11000

RESULT 11
SPBC1703/c
LOCUS SPBC1703 32799 bp DNA PLN 24-JAN-2000
DEFINITION S.pombe chromosome II cosmid c1703.
ACCESSION AL136536
VERSION AL136536.1 GI:6723914
KEYWORDS 5-formyltetrahydrofolate cyclo-ligase; ARID DNA binding domain; Armadillo/beta-catenin-like repeats; ATP citrate lyase; CoA-ligases; coiled-coil; conserved hypothetical; DNA mismatch repair protein; dna topoisomerase I; F-box domain; mitochondrial phosphate carrier protein; MLH1 homolog; pop4 protein homolog; Ras family; RIO1/ZK632.3/MJ0444 family; Sec1 family; top1; transfer RNA; tRNA Pro anticodon AGG; ubiquitin carboxyl-terminal hydrolase; vacuolar protein sorting protein; ypt1; ypt1-related protein 1.
SOURCE fission yeast.
ORGANISM Schizosaccharomyces pombe
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes; Schizosaccharomycetales; Schizosaccharomycetaceae; Schizosaccharomyces.
REFERENCE 1 (bases 1 to 32799)
AUTHORS McDougall,R.C., Rajandream,M.A., Barrell,B.G., Cadieu,E., Lelaure,V. and Galibert,F.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-2000) European Schizosaccharomyces genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and Laboratoire de Recombinaisons Genetiques, 2 avenue du Professeur Leon Bernard, Faculte de Medecine, F-35043 Rennes cedex, France
COMMENT Notes:
Details of yeast sequencing at the Sanger Centre are available on

```
            the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/S_pombe/)
            During 1995 to 1996 about 66% of S. pombe chromosome 1 was
            sequenced by the Sanger Centre.  The sequencing of the S. pombe
            genome is now being continued with funding from The European
            Commission. Fourteen European sequencing laboratories, including
            the Sanger Centre, are participating in the project.
            Protein coding regions (CDS) have been predicted with the help of
            computer analysis using the Genefinder program in PomBase (an ACEDB
            database) with additional predictions for the branch-acceptor sites
            supplied by the program Sp3splice. CAUTION: It is possible that for
            any individual CDS we may have underestimated or overestimated the
            number of introns/exons or we may not have chosen the correct
            splice donor/acceptor sites.
            CDS are numbered using the following system eg SPBC25H2.01c. SP (S.
            pombe), B (chromosome 2), c25H2 (cosmid name), .01 (first CDS), c
            (complementary strand).
            The more significant matches with motifs in the PROSITE database
            are also included but some of these may be fortuitous.
            The length in codons is given for each CDS.
            IMPORTANT: This sequence MAY NOT be the entire insert of the
            sequenced clone.  It may be shorter because we only sequence
            overlapping sections once, or longer, because we arrange for a
            small overlap between neighbouring submissions.
            Cosmid c1703 is overlapped at the 3' end by cosmid c2A9, EMBL entry
            SPBC2A9, accession number AL049495.
FEATURES            Location/Qualifiers
     source         1. .32799
                    /organism="Schizosaccharomyces pombe"
                    /strain="972h-"
                    /db_xref="taxon:4896"
                    /chromosome="II"
                    /clone="cosmid c1703"
                    /map="IIR"
     gene           complement(1. .721)
                    /gene="SPBC1703.01c"
     CDS            complement(join(1. .144,217. .721))
                    /partial
                    /gene="SPBC1703.01c"
                    /note="SPBC1703.01c, len:216, SIMILARITY:Homo sapiens,
                    CAB39167, pop4 protein., (220 aa), fasta scores: opt: 329,
                    E():3e-15, (36.9% identity in 160 aa)"
                    /codon_start=1
                    /label=SPBC1703.01c
                    /product="putative human pop4 protein homolog"
                    /protein_id="CAB66445.1"
                    /db_xref="GI:6723915"
                    /translation="MADPLYSSLKTVNSTSKVTPESLVFQTKILTPEEAKNVINEKIA
                    FKPLLLVPSLNLEKQGDRKESVASKKRKKLSSKEKKKLQLNVVPKIADYSQFKHLHSM
                    WCSYILEVIAGCTGESLMAKLAKAEYQGAYMQVLRSKSTTRVGLEGICIHESKHMLSL
                    ITKENRVVRVPKQDSVMKVIVDVPQRKLVFELYTQHLLLRAGDRSNKRFKSKNTID"
     misc_feature   complement(145. .157)
                    /gene="SPBC1703.01c"
                    /note="ctaattcgtctag, splice branch and acceptor"
     misc_feature   complement(211. .216)
                    /gene="SPBC1703.01c"
                    /note="gtaagt, splice donor sequence"
     gene           1398. .3793
                    /gene="SPBC1703.02"
     CDS            join(1398. .1510,1564. .3793)
                    /gene="SPBC1703.02"
                    /note="SPBC1703.02, len:780, SIMILARITY:Saccharomyces
                    cerevisiae, YMM7_YEAST, hypothetical 65.2 kd protein in
                    cox14-hmgs intergenic region., (581 aa), fasta scores:
                    opt: 363, E():5e-16, (26.1% identity in 528 aa)"
                    /codon_start=1
                    /label=SPBC1703.02
                    /product="hypothetical protein with ARID DNA-binding
                    domain"
                    /protein_id="CAB66446.1"
                    /db_xref="GI:6723916"
                    /translation="MSQVIQGAVEEPVYSNKSLTNENDSFLSLIESFSQERGVPIDIN
                    PKIGRKPILLYELYKKVIKRGGYDAVSATEDGWTNIAEEFNQSDPARSAGILQNVYFK
                    YLISWEIHDHWHLLLPPPSTLELSENRQNVLERVKVFNSCSPPKPTNPITIIDNDKTH
                    SFKKFYKSLVKEDEENGIEKKINEAMNSSASQPLPNSTNFASTTFPSVPFHPLPVDSG
                    LQKYIDRNSNLTPPALGAGVPGPPLLVRVALALKSQQDKEVDWAIGHLIKISFERHQE
                    FKLERLPSLAECLVYSLGFQVTKAKQVSDLTDISLCMDRAIGIALVLRNSVLSVENAK
                    HVAQTKLVISVLEASIRCAKTFNNLEFLHYCLDISEMISSYLHVEDEKNVFYLALCEF
                    LNSSDYSILIATLRTLARLALNDRNNRLLQDLKSNVIANIIALVETDNEEIVAAALDF
                    LYQYTTYRTNTSNLLASQEAWMLVDQLIRLMMYQAQERFMTVTIENSESNPAVKHEAT
                    SSISLPMEELQQLVAMREPERSVKWMRCCFEPSSDDYVLQTDIWQLYCSDMERAQGPN
                    VMGISPADFIKVSSHAFYNARAMTVSTPQLPVEYVINGIKRRKFPTSVNGMRYQPCRW
                    CLDSGKECGELLLGTPLLHSHLQEMHIFPQILETGKCRWSDCKYEIQRLTPASELSHY
                    QLLSHIVTHLHDDSLETLVEGRKLSPSREFRIPLLLTAVDDQGDATGIALTATLVLRN
                    LVRSKQGKMLFSAIESRVLEVSSLNPAVAPYVSEMLLGQI"
     misc_feature   1446. .1781
                    /gene="SPBC1703.02"
                    /note="Pfam match to entry PF01388 ARID, ARID DNA binding
                    domain, score 30.50, E-value 2.1e-06"
     misc_feature   1511. .1516
                    /gene="SPBC1703.02"
                    /note="gtaggt, splice donor sequence"
     misc_feature   1550. .1563
                    /gene="SPBC1703.02"
                    /note="ctaatccttttttag, splice branch and acceptor"
     gene           complement(4094. .6173)
                    /gene="SPBC1703.03c"
     CDS            complement(join(4094. .5959,6045. .6173))
                    /gene="SPBC1703.03c"
                    /note="SPBC1703.03c, len:664, SIMILARITY:Saccharomyces
                    cerevisiae, Q07395, orf ydl063c., (620 aa), fasta scores:
                    opt: 241, E():8.3e-08, (22.3% identity in 683 aa)"
                    /codon_start=1
                    /label=SPBC1703.03c
                    /product="hypothetical Armadillo/beta-catenin domain
                    protein"
                    /protein_id="CAB66447.1"
                    /db_xref="GI:6723917"
                    /translation="MGRANGKRRNHRKNRMNPIQKKGFLEQEIPLNNSEAALPLLSRL
                    DNPEIKERSWACSAISNIIASCTEGRLYLLKNGLVSKLIDRISDDSVEVVVEATGALR
                    NLAIEEGYSICMDMYRKNVLAPLQIWQTKIVQTLDEATEGKNVLETYEDRSTFSCLCA
                    IAENISSLLVNLGETTSQVIKVLNQKNTLVFLSRLLIPEAKIPQSVQEMALQAFFTLT
                    DDNDDGIITWIQNSSDFALKTINQTYLYFSYPSCLVRVYSIGIIYNIYQSGFVNKKME
                    SVVKGISLFDSFIPEALPILSDLLPSEENYRNLVRQVYDKDTYFKTKKDDLTLSSELL
                    VIPATLELISSMTSLLQSLADGTDELEEQEDSLMEDEDLSYMDDMSNVVNEDENLIID
                    EIPSNTPQKGNFKLVEYMLDHVLPKVITYCVVAFEFSSEEIPSNLSNYFQEVGDRTIE
                    CLNNISWSCNSVFVESSEAFTRWKLSAGKILQWIFQTIFLRLGLGVWPFSSEGFTTAC
                    SLLWSVSKPFPAEIQVLSVDDISTLILFSSTHGSLEAQSRLLGAFGSLGRCGNIQINQ
                    LLGQTLMSCVIASDPNPLLAVEALNAIFDVYGDKSYPYDAPVFKGSGYLSQLSEALPR
                    LKNMVKKIDRRREKHLRFRAEEALENLESFIDYKHAEYAS"
     misc_feature   complement(5856. .5978)
                    /gene="SPBC1703.03c"
                    /note="Pfam match to entry PF00514 Armadillo_seg,
                    Armadillo/beta-catenin-like repeats, score 30.60, E-value
                    3.6e-05"
     misc_feature   complement(5960. .5976)
                    /gene="SPBC1703.03c"
                    /note="ctaactcttctaattag, splice branch and acceptor"
     misc_feature   complement(6039. .6044)
                    /gene="SPBC1703.03c"
                    /note="gtatgt, splice donor sequence"
     gene           6367. .8661
                    /gene="SPBC1703.04"
     CDS            join(6367. .6438,6567. .6704,6770. .6942,6990. .8661)
                    /gene="SPBC1703.04"
                    /note="SPBC1703.04, len:684, SIMILARITY:Arabidopsis
                    thaliana, Q9ZRV4, MLH1 PROTEIN, (737 aa), fasta scores:
                    opt: 1138, E():0, (33.7% identity in 736 aa);
                    non-consensus branch site in second intron"
                    /codon_start=1
                    /label=SPBC1703.04
                    /product="putative DNA mismatch repair protein, MLH1
                    homolog"
                    /protein_id="CAB66448.1"
                    /db_xref="GI:6723918"
                    /translation="MDVNSRAKIRPLDQLVINKIAAGEIIERPENAIKELIENSLDAG
```

```
                     STSIDVLLKDGGLKLLQITDNGSGIQYDDLPYLCQRFSTSKIDNFNDLQHLQTFGFRG
                     EALASISHVAKVTVVTKLSSDIHAWKAFYVDGALAPISPGMSPAPQPCAGKQGTVITA
                     EDLFYNVRSRKSALKNGSEEFRRIMILVQKYAIHNDQVSFNCKKVGDTVASLSLSSRL
                     SKADKIRHIYGPRVASHLRDFSLGEGQSSIVGFSANGFISNADFQDKKSNLILFINNR
                     LVESVELRHALEETYAKYLHKGASYFVYLSLNMSPEQLDVNVHPSKRIVHFLYDQEIA
                     TSICDKLGEILERTDTERSYPLQAMIPSISNTKNAESSSQKAVRTYENYLVRTDPRER
                     SIKSMLSDNFLQRSSNNYDNEIIEKVDSANSNKNATNDIKDLQTEEIVEEGNSIDLES
                     IKSLQKQVINSMHVLATNILTEHKYVGLVCPTRRIAAVQHNIGLYYVDYGKLSYHLFY
                     QICLTEFGNYGEFVLETPLSISDLFEIVNGDEDKSESEKFTRLLVSRRDMLKDYFSIS
                     VTSGGLLTAVPMLSPKYHPPFEQLPLLISSLTPKFFDWLDEKSCLNGIMKAIAKFYVP
                     LPLSYEESDVKSIRSLESCLEDYLFPEFRRRVICPKKVFEEKCIYQITSLPRLYNVFE
                     RC"
     misc_feature    join(6391. .6438,6567. .6704,6770. .6942,6990. .7617)
                     /gene="SPBC1703.04"
                     /note="Match to PF01119 DNA_mis_repair, DNA mismatch
                     repair protein Score 307.02"
     misc_feature    6439. .6444
                     /gene="SPBC1703.04"
                     /note="gtaagc, splice donor sequence"
     misc_feature    6554. .6566
                     /gene="SPBC1703.04"
                     /note="ctaactaattaag, splice branch and acceptor"
     misc_feature    6705. .6710
                     /gene="SPBC1703.04"
                     /note="gtaagt, splice donor sequence"

  Query Match             1.3%;  Score 20;  DB 15;  Length 32799;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      90 cgacacggttgatatctcat 109
           ||||||||||||||||||||
Db    3318 CGACACGGTTGATATCTCAT 3299


RESULT  12
SPAC22A12/c
LOCUS       SPAC22A12    32814 bp    DNA             PLN       04-SEP-2000
DEFINITION  S.pombe chromosome I cosmid c22A12.
ACCESSION   Z99295
VERSION     Z99295.1  GI:2414594
KEYWORDS    40s ribosomal protein s15a; aminoalcoholphosphotransferase; atp
            citrate lyase; bip; dihydrofolate reductase; dihydroxyacetone
            kinase; dolichyl-phosphate-mannose--protein mannosyl transferase;
            endoplasmic reticulum; glucose regulated protein; heat shock
            protein 70 family; Homol D box; oxidoreducatase; phosphatidyl
            synthase; pre-mrna splicing factor; rps15a-2; short chain
            dehydrogenase; sorbitol utilization; transcription elongation
            factor.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 32814)
  AUTHORS   Pearson,D., Churcher,C.M., Barrell,B.G., Rajandream,M.A. and
            Wood,V.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-SEP-1997) Schizosaccharomyces pombe chromosome I
            sequencing project, Sanger Centre, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA E-mail: barrell@sanger.ac.uk
COMMENT     Notes:
            Details of yeast sequencing at the Sanger Centre are available on
            the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/S_pombe/)
            Protein coding regions (CDS) have been predicted with the help of
            computer analysis using the Genefinder program in PomBase (an ACEDB
            database) with additional predictions for the branch-acceptor sites
            supplied by the program Sp3splice. CAUTION: It is possible that for
            any individual CDS we may have underestimated or overestimated the
            number of introns/exons or we may not have chosen the correct
            splice donor/acceptor sites.
            CDS are numbered using the following system eg SPAC5H10.01c. SP (S.
            pombe), A (chromosome 1), c5H10 (cosmid name), .01 (first CDS), c
            (complementary strand).
            The more significant matches with motifs in the PROSITE database
            are also included but some of these may be fortuitous.
            The length in codons is given for each CDS.
            IMPORTANT: This sequence MAY NOT be the entire insert of the
            sequenced clone.  It may be shorter because we only sequence
            overlapping sections once, or longer, because we arrange for a
            small overlap between neighbouring submissions.
            Cosmid c22A12 is overlapped at the 5' end by cosmid SPAC56F8,
            EM:Z69728 and at the 3' end by cosmid SPAC4C5, EM:Z98560.
FEATURES            Location/Qualifiers
     source         1. .32814
                    /organism="Schizosaccharomyces pombe"
                    /strain="972h-"
                    /db_xref="taxon:4896"
                    /chromosome="I"
                    /clone="cosmid c22A12"
                    /map="IL"
     misc_feature   1. .218
                    /note="nominal overlap with cosmid SPAC56F8, EM:Z69728 S.
                    pombe chromosome 1"
     CDS            complement(1. .317)
                    /partial
                    /gene="SPAC22A12.01c"
                    /codon_start=1
                    /label=SPAC22A12.01c
                    /product="DNA repair protein for interstrand crosslinks;
                    by similarity to yeast pos2"
                    /protein_id="CAB16571.2"
                    /db_xref="GI:6855449"
                    /db_xref="SWISS-PROT:Q10264"
                    /translation="MSKRKSSSILDFFKKSNNGKNMWGNEAVSVEKTDSTTKPSTVKN
                    AISSSQIKNEKFSQEVLLQFNDPSFTSEVGESPQWQSFGDANILEPLKNELVKNEEST
                    MSE"
     gene           complement(1. .317)
                    /gene="SPAC22A12.01c"
                    /note="SPAC56F8.17c"
     gene           complement(1151. .1525)
                    /gene="SPAC22A12.02c"
     CDS            complement(1151. .1525)
                    /gene="SPAC22A12.02c"
                    /note="SPAC22A12.02c, len:124aa"
                    /codon_start=1
                    /label=SPAC22A12.02c
                    /product="very hypothetical protein"
                    /protein_id="CAB16572.1"
                    /db_xref="GI:2414595"
                    /db_xref="SPTREMBL:O13894"
                    /translation="MEQFEIIRESDADPYYFNKNLRLEECEIDGFTSIYKAPPPYSFV
                    AGLSKNTSSNSFVIHKCLPPLQRNNTIRKRKRNPFLADRPYCINSEDMPLSDDRDMLK
                    ERHVKSRKLLLKEKGNRFYRRS"
     gene           complement(1744. .2877)
                    /gene="SPAC22A12.03c"
     CDS            complement(1744. .2877)
                    /gene="SPAC22A12.03c"
                    /note="SPAC22A12.03c, len:377aa; similarity: some to C.
                    int estinalis, P91586, COS41.8, (409aa), fasta scores,
                    opt:240, E():1.6e-05, (23.3% identity in 258 aa overlap)"
                    /codon_start=1
                    /label=SPAC22A12.03c
                    /product="putative COP complex subunit; contains PCI
                    domain"
                    /protein_id="CAB16573.1"
                    /db_xref="GI:2414596"
                    /db_xref="SPTREMBL:O13895"
                    /translation="MEEVVHYFLEGNMPVAQFREALALHYTNEKELFEQAKRCLNICC
                    GSNNFAKRNDVLFSLLDVAVSISSLELRKELISELYVPVQSLEEAPSEYLVSCCLQLA
                    TIYEAEQNFELLCSSLEAVEKHGHFENDLEQLLLLRIRLGDAYLKLGKAEKAILTVRT
                    SIPLAFKVSNDQLLMELQLCNARALDETGQFLEAAKCYYRVLQYKVPGNELIYRENLC
                    SVAQCLLLAIPSPIVLQFLQEISLQPSVREIPFYSLVEKYLKRKFIGKEDGAFLLPFL
                    LPHQVIHMNRLIEDGRHFLETNILFLSEFFEVSSTSILAKHFKLSEEQVDTVVADMVI
                    QERLNASIDQCEGYITFLPEYGKANNLPNYVNKIATVLQHYGS"
```

```
misc_feature    complement(1795. .2031)
                /gene="SPAC22A12.03c"
                /note="Match to PF01399 PCI, PCI domain Score 88.36"
gene            complement(3769. .4161)
                /gene="SPAC22A12.04c"
                /note="rps15a-1; rps22-1"
CDS             complement(3769. .4161)
                /gene="SPAC22A12.04c"
                /note="SPAC22A12.04c, len:130aa"
                /codon_start=1
                /product="40s ribosomal protein S15A/S22A"
                /protein_id="CAB16574.1"
                /db_xref="GI:2414597"
                /db_xref="SWISS-PROT:O14469"
                /translation="MVRQSVLADCLNNIVNAERRGRRQVLIRPSSKVIVKFLTVMQKH
                GYIDEFTEIDDHRSGKIVIQLNGRINKCGVISPRFNVKLKDIEKWVNQLLPSRQVGVI
                VLTTSRGIMSHNEARAKDAGGKILGFFY"
misc_feature    complement(3772. .4149)
                /gene="SPAC22A12.04c"
                /note="Pfam match to entry PF00410 S8, Ribosomal protein
                S8"
misc_feature    complement(3811. .3864)
                /gene="SPAC22A12.04c"
                /note="PS00053 Ribosomal protein S8 signature"
promoter        complement(4232. .4239)
                /note="Homol D box"
gene            4648. .4977
                /gene="SPAC22A12.05"
                /note="rpc11"
CDS             4648. .4977
                /gene="SPAC22A12.05"
                /note="SPAC22A12.05, len:109"
                /codon_start=1
                /label=rpc11
                /product="dna-directed rna polymerase iii 12.5 kd
                polypeptide"
                /protein_id="CAB16575.1"
                /db_xref="GI:2414598"
                /db_xref="SWISS-PROT:O13896"
                /translation="MQFCPTCGNHLIVAVDEEGRNAFDCRTCPYHFPISTFLYSRHEF
                AQKEVDDVLGGEEAFESNQQTEVTCENTKCDNNRAYFFQLQIRSADEPMSTFYRCTKC
                KFQWREN"
gene            complement(5461. .6750)
                /gene="SPAC22A12.06c"
CDS             complement(5461. .6750)
                /gene="SPAC22A12.06c"
                /note="SPAC22A12.06c, len:429, Low
                SIMILARITY:Schizosaccharomyces pombe, DYR_SCHPO,
                dihydrofolate reductase, (461 aa), fasta scores: opt: 438,
                E():3.6e-22, (23.3% identity in 455 aa)"
                /codon_start=1
                /label=SPAC22A12.06c
                /product="similar to S. pombe dihydrofolate reductase
                Dfr1p"
                /protein_id="CAB16576.1"
                /db_xref="GI:2414599"
                /db_xref="SPTREMBL:O13897"
                /translation="MKSATKSKILCIHGYAESGELFSVKLRALRERMADSVDFYFPTG
                PIELDKAKDELNGSGFDALSTVFSSSPASHRRGWWRINEYADTKQLEPTKAFEYLASY
                IKEHGPFDGILGFSQGTNLAANLAALVTIPKYQEYFSQPPFRFALFFSGYFRPLLMDG
                AVHATKLDLPTLHLLGKYDTVLSTETSTTLVRACKDAQVLFHPAAHQIPAPHAYVEPA
                ADFIDFFSREDWPIISKHISLIVPTKKVNTTSAQTSLNNVEHDLISKIMSRSFKGGIN
                VVVSDLQMFNEYKRIFGPKVLSLVVTDGQEPDTYSENVHYTPNFKGALAYLEAYQNSI
                GRIYVIGDKKLLTLGMLCRCTKRIIAITDAEDKFISQSSAQQVSGSLPFLEHSKEWLK
                AKSSQIRQWTGQSRLKHMSQDNSGEPVTMKLQMWERL"
gene            complement(8347. .11028)
                /gene="SPAC22A12.07c"
CDS             complement(8347. .11028)
                /gene="SPAC22A12.07c"
                /note="SPAC22A12.07c, len:893aa, similarity: to YDL095W,
                PM T1_YEAST, P33775, dolichyl-phosphate-mannose--protein
                manno syltransferase 1, (817aa), fasta scores, opt:1505,
                E():0, ( 33.2% identity in 747 aa overlap)
                SPAC22A12.07c, len:893, SIMILARITY:Candida albicans,
                PMT1_CANAL, dolichyl-phosphate-mannose--protein
                mannosyltransferase 1(ec 2.4.1.109)., (877 aa), fasta
                scores: opt: 1558, E():0, (33.9% identity in 832 aa)"
                /codon_start=1
                /label=SPAC22A12.07c
                /product="dolichyl-phosphate-mannose--protein
                mannosyltransferase 1(ec 2.4.1.109)."
                /protein_id="CAB16577.1"
                /db_xref="GI:2414600"
                /db_xref="SWISS-PROT:O13898"
                /translation="MDKQSTFQDPKEKHRIQRDVKLSRPRKRFSFLDYVVVIFLTVVA
                FCVRAQRLMNPAKVVFEELRYYNYAVDYVNNKLLMDVYPPLGKLLFSLVAALTGNKYE
                LNTLDEPGQQYPFTDVAYSMRLFTCLLGSLLVPLMYGTVYFPTKSKTAASLAALFVIF
                DNGLITMSRYIMIEIPALYFMSLTAFYWSVYEAQQKRPFSLRWHTSLLSTGVALGLAL
                STKLSAMFTFGWLLILAAFHLWNLLGDLSVPMYRIVKHLFSYIFYLIGVPITVYLAVF
                AVHSHIAYKASVADAFLPPEHRHALAGNRFDDQFADVAYGSLVTIRNAIPEHGYLHSS
                ELLYPEGTEQQIISLVDEPNQNALWIIEHEHSQDNNRSNIELLKDGSVVRLRHVMTGR

  Query Match             1.3%;  Score 20;  DB 15;  Length 32814;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    731 agaagctgcgaaatgttact 750
          ||||||||||||||||||||
Db   2302 AGAAGCTGCGAAATGTTACT 2283


RESULT  13
AC003975/c
LOCUS       AC003975    41025 bp    DNA             PRI       02-JAN-1998
DEFINITION  Human Cosmid g5129s420 from 7q31.3, complete sequence.
ACCESSION   AC003975
VERSION     AC003975.1  GI:2739390
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 41025)
  AUTHORS   Iadonato,S.P., Yu,J., Wong,G.K.-S., Magness,C.L., Green,E.D.,
            Green,P. and Olson,M.V.
  TITLE     Large-scale MCD Mapping and Sequencing of Human Chromosome 7
  JOURNAL   Unpublished (1996)
  REMARK    University of Washington Human Genome Center
              Box 352145 Seattle, WA 98195
            Contact: Shawn Iadonato (iadonato@u.washington.edu)
REFERENCE   2  (bases 1 to 41025)
  AUTHORS   Iadonato,S.P., Yu,J., Wong,G.K.-S., Magness,C.L., Green,E.D.,
            Green,P. and Olson,M.V.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JAN-1998) Human Genome Center, University of
            Washington, Box 352145, Seattle, WA 98195, USA
  REMARK    University of Washington Human Genome Center
              Box 352145 Seattle, WA 98195
              Contact: Shawn Iadonato (iadonato@u.washington.edu)
COMMENT     Overlapping Sequences:
               5': UWGC:(mapping in progress)
               3': UWGC:g5129z079
              --------------------
            Sequence Quality Assessment:
             This entry has been annotated with sequence quality
             estimates computed by the Phrap assembly program.
             All manually edited bases have been reduced to quality zero.
             Quality levels above 40 are expected to have less than
             1 error in 10,000 bp.

             Pct of Consensus above quality 40: 98.7%
             Number of manually edited bases:    2

             Double stranded (DS) coverage:  75.1%
             DS or two chemistry coverage:  100.0%
             Single stranded regions:           0
```

```
Quality  Cumulative Percentage of Consensus
-------  ------------------------------
    90   xxx (15.2%)
    80+  xxxxxxxxxx (52.1%)
    70+  xxxxxxxxxxxxxxx (84.1%)
    60+  xxxxxxxxxxxxxxxxx (92.9%)
    50+  xxxxxxxxxxxxxxxxx (96.6%)
    40+  xxxxxxxxxxxxxxxxxx (98.7%)
    30+  xxxxxxxxxxxxxxxxxx (99.8%)
    20+  xxxxxxxxxxxxxxxxxx (100.0%)
    10+  xxxxxxxxxxxxxxxxxx (100.0%)
    00+  xxxxxxxxxxxxxxxxxx (100.0%)

Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.
```

Sequence Validation:
This sequence has been validated by Multiple Complete Digest Mapping. Comparison of the experimentally derived map digest fragments with sequence-predicted fragments is given below. Small fragments below a variable cutoff (approximately 400-600bp) are not mapped and hence do not appear in the table. There are no significant remaining descrepancies between the experimental and predicted values. Uniquely ordered fragment groups are separated by dashed lines.

EcoRI		HindIII		NsiI	
Map	Seq	Map	Seq	Map	Seq
1391.43	1396.00	798.14	781.00	1441.14	1454.00
662.43	646.00	1901.71	1909.00	3681.00	3734.00
3239.62	3253.00	1194.88	1199.00	2105.75	2122.00
2971.38	3011.00	9182.75	9198.00	1103.62	1109.00
1737.88	1755.00	8076.38	8057.00	4724.29	4681.00
12631.25	12699.00	1984.42	1989.00	5616.08	5621.00
7091.09	7178.00	7033.18	7078.00	1630.74	1625.00
2876.31	2856.00	1316.12	1317.00	3683.24	3716.00
5061.81	5094.00	1112.53	1102.00	4722.89	4741.00
829.37	825.00	6496.97	6491.00	1831.07	1833.00
1007.39	1007.00			1551.92	1575.00
				1667.89	1670.00
				926.54	923.00

```
FEATURES             Location/Qualifiers
     source          1. .41025
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q31.3"
                     /clone="NHGRI:yWSS5129 or CEPH:744C02"
                     /sub_clone="UWGC:g5129s420"
                     /clone_lib="CEPH MegaYACs"
     repeat_region   3405. .3738
                     /rpt_family="L1"
     repeat_region   complement(5381. .5739)
                     /rpt_family="THE1"
     repeat_region   complement(7338. .7776)
                     /rpt_family="L1"
     repeat_region   12864. .13215
                     /rpt_family="THE1"
     repeat_region   13230. .14227
                     /rpt_family="MSTAR"
     repeat_region   14354. .15323
                     /rpt_family="LTR5"
     repeat_region   17865. .17976
                     /rpt_family="MLT2B2"
     repeat_region   complement(19181. .19540)
                     /rpt_family="THE1"
     repeat_region   20230. .20507
                     /rpt_family="ALU"
     repeat_region   22545. .22818
                     /rpt_family="ALU"
     repeat_region   26258. .26590
                     /rpt_family="THE1"
     repeat_region   28100. .28229
                     /rpt_family="ALU"
     repeat_region   complement(28314. .28602)
                     /rpt_family="ALU"
     repeat_region   complement(28624. .29217)
                     /rpt_family="L1"
     repeat_region   29244. .30172
                     /rpt_family="L1"
     repeat_region   36764. .37542
                     /rpt_family="L1"
     repeat_region   37551. .37846
                     /rpt_family="ALU"
     repeat_region   38064. .38152
                     /rpt_family="L1"
     repeat_region   complement(39840. .40132)
                     /rpt_family=">MER42"
BASE COUNT    13471 a   7112 c   7207 g  13235 t
ORIGIN


  Query Match             1.3%;  Score 20;  DB 85;  Length 41025;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      30 agttctgcttttaatttttta 49
           ||||||||||||||||||||
Db  18250 AGTTCTGCTTTTAATTTTTA 18231


RESULT  14
AL355869/c
LOCUS       AL355869   103734 bp    DNA             HTG       09-FEB-2001
DEFINITION  Homo sapiens chromosome 1 clone RP5-888O16, *** SEQUENCING IN
            PROGRESS ***, in unordered pieces.
ACCESSION   AL355869
VERSION     AL355869.8  GI:12733618
KEYWORDS    HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103734)
  AUTHORS   Frankland,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 9, 2001 this sequence version replaced gi:12043422.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: dJ888O16
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
```

```
           Consensus quality: 103438 bases at least Q40
           Consensus quality: 103734 bases at least Q30
           Consensus quality: 103734 bases at least Q20
           Insert size: 103734; sum-of-contigs
           Insert size: 124020; 32.2% error; agarose-fp
           Quality coverage: 5.33x in Q20 bases; sum-of-contigs Quality
           coverage: 4.46x in Q20 bases; agarose-fp
           ---------------
           * NOTE: This is a 'working draft' sequence.
           * This record will be updated with the finished sequence
           * as soon as it is available and the accession number will
           * be preserved.
FEATURES             Location/Qualifiers
     source          1. .103734
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP5-888016"
                     /clone_lib="RPCI-5"
     misc_feature    1. .103734
                     /note="assembly_fragment:00084"
BASE COUNT    33135 a  17664 c  18416 g  34519 t
ORIGIN


  Query Match             1.3%;  Score 20;  DB 79;  Length 103734;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 tgggtgattttctgtacagt 32
           ||||||||||||||||||||
Db   63426 TGGGTGATTTTCTGTACAGT 63407


RESULT  15
AL137126_3/c
WPCOMMENT
  Sequence split into 5 fragments  LOCUS AL137126 Accession AL137126
     Fragment Name      Begin     End
     AL137126_0             1    110000
     AL137126_1        100001    210000
     AL137126_2        200001    310000
     AL137126_3        300001    410000
     AL137126_4        400001    413028
Continuation (4 of 5) of AL137126 from base 300001 (AL137126 Homo sapiens chromosome 9 c


  Query Match             1.3%;  Score 20;  DB 78;  Length 110000;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 tgggtgattttctgtacagt 32
           ||||||||||||||||||||
Db   84703 TGGGTGATTTTCTGTACAGT 84684


Search completed: October 30, 2001, 13:13:41
Job time: 4468 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2001, 12:07:53 ; Search time 119.82 Seconds
(without alignments)
8049.215 Million cell updates/sec

Title: US-09-619-049-854
Perfect score: 1536
Sequence: 1 atgaggcttccttgggtgat.....ggactctaatgatggattct 1536

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 730101 seqs, 313950809 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N_Geneseq_0601:*
1: /SIDS1/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	21	1.4	400	21	AAA75359	Fragment of IRES r
c 2	20	1.3	683	19	AAV52471	Streptococcus pneu
3	19	1.2	262	21	AAA69926	Human ovarian carc
c 4	19	1.2	282	21	AAA78188	cDNA encoding huma
c 5	19	1.2	307	21	AAA69935	Human ovarian carc
c 6	19	1.2	977	18	AAT80189	Human collagen 1A1
c 7	19	1.2	2192	21	AAC77648	Human cancer assoc
8	19	1.2	3871	8	AAN71302	HSV-1 gB and surro
9	19	1.2	3996	6	AAN50516	Sequence encoding
c 10	18	1.2	430	16	AAT20801	Human gene signatu
c 11	18	1.2	528	21	AAZ29568	Rat cDNA clone rC1
c 12	18	1.2	546	21	AAZ34777	Rice sucrose phosp
13	18	1.2	930	21	AAC66433	Human secreted pro
14	18	1.2	1149	22	AAC86928	DNA encoding nucle
15	18	1.2	1689	21	AAA70248	Plasmodium falcipa
c 16	18	1.2	1988	18	AAT72710	Mouse inhibitor of
17	18	1.2	2126	20	AAZ00448	Human secreted pro
c 18	18	1.2	1082138	21	AAF22305	Arabidopsis thalia
c 19	18	1.2	1830121	17	AAT42063	Haemophilus influe
20	17	1.1	252	20	AAV89430	EST clone CO908.
21	17	1.1	283	20	AAV88657	EST clone GZ420.
c 22	17	1.1	284	21	AAA31552	Plant microsatelli
23	17	1.1	303	17	AAT09210	Virulence factor s
c 24	17	1.1	343	21	AAA31442	Plant microsatelli
25	17	1.1	360	21	AAA81870	N. meningitidis pa
26	17	1.1	399	22	AAF64541	Novel human polynu
c 27	17	1.1	403	21	AAA31949	Plant microsatelli
c 28	17	1.1	448	21	AAC24250	Human secreted pro
c 29	17	1.1	499	21	AAA31838	Plant microsatelli
30	17	1.1	637	10	AAN90220	Malaria-specific D
c 31	17	1.1	744	21	AAF14271	Aspergillus oryzae
c 32	17	1.1	888	21	AAC42857	Arabidopsis thalia
33	17	1.1	1155	22	AAF28259	Murine TANGO 300 o
34	17	1.1	1212	21	AAZ51995	Salmonella Pathoge
35	17	1.1	1400	22	AAF28258	Murine TANGO 300 D
c 36	17	1.1	1445	21	AAZ34781	Wheat sucrose phos
37	17	1.1	1791	21	AAZ46369	Haemorrhagic enter
c 38	17	1.1	3521	22	AAC84583	Rice magnesium che
39	17	1.1	5139	21	AAA70139	Plasmodium falcipa
c 40	17	1.1	5746	17	AAT09225	Partial sequence o
c 41	17	1.1	7769	18	AAV74525	Staphylococcus aur
42	17	1.1	8457	21	AAZ51984	Genomic sequence-1
c 43	17	1.1	9278	16	AAQ76124	Human MDC genomic
c 44	17	1.1	11272	20	AAX89842	Rat cubilin protei
c 45	17	1.1	26270	21	AAZ46355	Haemorrhagic enter
46	17	1.1	65792	22	AAF28544	Genomic fragment #
c 47	17	1.1	92584	21	AAF22288	BAC containing rep
48	17	1.1	1830121	17	AAT42063	Haemophilus influe
c 49	16	1.0	24	21	AAZ56479	Vascular endotheli
c 50	16	1.0	153	21	AAC54565	Arabidopsis thalia
c 51	16	1.0	177	21	AAC25196	Human secreted pro
c 52	16	1.0	259	21	AAA00900	Human colon cancer
c 53	16	1.0	265	21	AAC09614	Human secreted pro
54	16	1.0	272	20	AAZ13221	Human gene express
c 55	16	1.0	314	20	AAX20145	Enterococcus faeca
56	16	1.0	395	21	AAC07628	Human secreted pro
c 57	16	1.0	423	20	AAX20144	Enterococcus faeca
58	16	1.0	424	16	AAT19917	Human gene signatu
c 59	16	1.0	450	20	AAX79205	Anti-HIV-1 gp120 a
c 60	16	1.0	456	21	AAC76749	Human ORFX ORF2304
61	16	1.0	483	21	AAC75455	Human ORFX ORF1010
c 62	16	1.0	490	19	AAV38809	Homo sapiens CESP
63	16	1.0	525	20	AAV88055	EST clone FR436.
c 64	16	1.0	547	22	AAF68147	Human lung tumour
65	16	1.0	571	21	AAF08239	Fusarium venenatum
c 66	16	1.0	609	22	AAF65595	Novel human polynu
67	16	1.0	646	20	AAX79022	Human secreted pro
c 68	16	1.0	722	21	AAZ97400	Human prostate can
c 69	16	1.0	745	20	AAX80505	cDNA of Apoptin-as
70	16	1.0	821	21	AAF15050	Trichoderma reesei
c 71	16	1.0	831	21	AAC36992	Arabidopsis thalia
72	16	1.0	917	21	AAC95082	Cat flea hindgut a
73	16	1.0	999	21	AAZ54302	Neisseria meningit
74	16	1.0	999	21	AAZ54303	Neisseria meningit
75	16	1.0	1016	21	AAC44334	Arabidopsis thalia
76	16	1.0	1020	20	AAZ12038	Neisseria gonorrho
77	16	1.0	1034	21	AAC38067	Arabidopsis thalia
78	16	1.0	1068	21	AAA58928	cDNA encoding a Ne
c 79	16	1.0	1074	21	AAA15956	Human protein clon
80	16	1.0	1110	16	AAT07301	Type I matrix meta
81	16	1.0	1120	22	AAC88073	Human FLEXHT-4 nuc
82	16	1.0	1141	21	AAZ95381	Human colon specif
c 83	16	1.0	1186	18	AAV74471	Staphylococcus aur
84	16	1.0	1220	21	AAC40496	Arabidopsis thalia

85	16	1.0	1226	15	AAQ70664	ScFv-BSRNase fusio
86	16	1.0	1235	15	AAQ70663	ScFv-BSRNase fusio
87	16	1.0	1259	15	AAQ70662	ScFv-BSRNase fusio
88	16	1.0	1260	15	AAQ70668	H17-Dip. Tox.-BSRN
89	16	1.0	1267	20	AAX97978	Human secreted pro
90	16	1.0	1272	15	AAQ70669	H17-Dip. Tox.-BSRN
91	16	1.0	1291	21	AAA26287	Human secreted pro
92	16	1.0	1292	21	AAA26331	Human secreted pro
93	16	1.0	1317	20	AAZ06368	Human receptor pro
94	16	1.0	1326	21	AAC45724	Arabidopsis thalia
c 95	16	1.0	1376	20	AAX06099	Human DnaJ-like pr
c 96	16	1.0	1386	14	AAQ49834	Anti-HIV-1 recombi
c 97	16	1.0	1475	21	AAC44799	Zea mays DNA fragm
98	16	1.0	1491	21	AAC45128	Arabidopsis thalia
c 99	16	1.0	1590	21	AAC61888	cDNA encoding a hu
100	16	1.0	1604	22	AAC84167	Human OX2R homolog
c 101	16	1.0	1608	22	AAC91988	Human T-bet coding
c 102	16	1.0	1619	21	AAA15966	Human protein clon
c 103	16	1.0	1621	21	AAZ64993	Membrane-bound pro
c 104	16	1.0	1621	22	AAF44139	Human PRO1080 (UNQ
c 105	16	1.0	1628	21	AAZ65326	Human secreted pro
106	16	1.0	1639	20	AAX32850	Mutant luciferase
107	16	1.0	1639	20	AAX32851	Mutant luciferase
108	16	1.0	1644	20	AAZ52913	Human prostate tum
c 109	16	1.0	1653	22	AAF67920	Corynebacterium gl
c 110	16	1.0	1666	22	AAF93807	Human cDNA encodin
c 111	16	1.0	1682	21	AAC98901	Human pancreatic c
c 112	16	1.0	1688	21	AAF18306	Lung cancer associ
c 113	16	1.0	1725	21	AAC76768	Human ORFX ORF2323
114	16	1.0	1744	15	AAQ63900	Acyl CoA synthetas
c 115	16	1.0	1746	21	AAC43162	Arabidopsis thalia
c 116	16	1.0	1836	21	AAC42921	Arabidopsis thalia
c 117	16	1.0	1984	21	AAC69571	Human secreted pro
118	16	1.0	2157	21	AAC54444	Arabidopsis thalia
119	16	1.0	2160	21	AAC52971	Arabidopsis thalia
c 120	16	1.0	2210	14	AAQ48431	Mycoplasma ferment
121	16	1.0	2218	19	AAX13957	H. pylori GHPO 702
c 122	16	1.0	2223	14	AAQ47809	Aspartokinase II g
c 123	16	1.0	2283	21	AAC74256	Human secreted pro
124	16	1.0	2427	22	AAF54905	Nucleotide sequenc
c 125	16	1.0	2445	21	AAZ45347	DNA encoding a Bac
c 126	16	1.0	2582	18	AAT90543	Maize male fertili
c 127	16	1.0	3276	21	AAC76770	Human ORFX ORF2325
c 128	16	1.0	3386	18	AAT84162	DNA encoding a lys
c 129	16	1.0	3386	19	AAV53504	DNA encoding a Sta
130	16	1.0	3410	9	AAN80617	Genomic clone enco
131	16	1.0	3411	9	AAN80644	Sequence of human
c 132	16	1.0	3533	20	AAV72867	Human lin-54 homol
133	16	1.0	3733	18	AAV74849	Staphylococcus aur
134	16	1.0	3751	20	AAZ42227	Human normal bladd
135	16	1.0	3793	21	AAA26458	Human secreted pro
136	16	1.0	3798	21	AAA26430	Human secreted pro
137	16	1.0	4065	21	AAF18226	Lung cancer associ
138	16	1.0	4138	15	AAQ55179	Human orphan recep
139	16	1.0	4610	21	AAA27896	Human GEF containi
140	16	1.0	4839	21	AAC47635	Arabidopsis thalia
c 141	16	1.0	4949	21	AAC76989	Human ORFX ORF2544
c 142	16	1.0	5263	21	AAA16627	Human secreted pro
c 143	16	1.0	5763	20	AAZ20606	Zea mays Rpg1-13 c
c 144	16	1.0	5872	18	AAT90544	Maize male fertili
c 145	16	1.0	7043	20	AAX13225	Enterococcus faeca
146	16	1.0	7156	20	AAX12966	Enterococcus faeca
147	16	1.0	7587	12	AAQ15133	Zucchini ACC synth
148	16	1.0	7587	19	AAV15703	Zucchini ACC synth
149	16	1.0	7587	22	AAF23620	Zucchini ACC synth
c 150	16	1.0	7715	19	AAV43042	Mus musculus Cgamm
151	16	1.0	7876	15	AAQ58707	E. coli multiple a
152	16	1.0	7876	19	AAV60383	wild type multiple
153	16	1.0	7878	21	AAC60488	E. coli MarR DNA.
154	16	1.0	7878	21	AAZ57660	Multiple antibioti
c 155	16	1.0	7959	21	AAA81510	N. meningitidis pa
156	16	1.0	7989	20	AAX13206	Enterococcus faeca
157	16	1.0	8395	20	AAX13154	Enterococcus faeca
158	16	1.0	8888	20	AAX02665	EP-892047 Seq ID 4
159	16	1.0	10470	18	AAV74331	Staphylococcus aur
160	16	1.0	11611	21	AAC86432	Wheat SSIII gene S
c 161	16	1.0	11864	19	AAV52194	Streptococcus pneu
162	16	1.0	20072	20	AAX13026	Enterococcus faeca
c 163	16	1.0	20072	20	AAX13026	Enterococcus faeca
c 164	16	1.0	20561	20	AAX13229	Enterococcus faeca
c 165	16	1.0	26778	21	AAA81477	N. meningitidis pa
c 166	16	1.0	28690	22	AAF57718	Human IL4Ralpha ge
c 167	16	1.0	30078	21	AAA81520	N. meningitidis pa
168	16	1.0	31096	18	AAV74370	Staphylococcus aur
c 169	16	1.0	35828	21	AAA29063	Murine TGF-beta bi
c 170	16	1.0	43576	21	AAA68609	Bacteriophage 96 c
171	16	1.0	44576	21	AAZ61522	Cosmid CVO14 conta
172	16	1.0	65921	21	AAZ89046	Human nibrin DNA.
173	16	1.0	96988	21	AAF22290	BAC containing rep
c 174	16	1.0	102634	21	AAA81464	N. meningitidis pa
175	16	1.0	349980	21	AAF21607	Neisseria meningit
c 176	16	1.0	349980	21	AAF21608	Neisseria meningit
c 177	16	1.0	349980	21	AAF21609	Neisseria meningit
178	16	1.0	349980	21	AAF21612	Neisseria meningit
179	16	1.0	1437668	21	AAA81490	N. meningitidis B
c 180	16	1.0	1437668	21	AAA81490	N. meningitidis B
181	15	1.0	15	22	AAF51816	IGF-I oligonucleot
c 182	15	1.0	17	21	AAA79914	Hepatitis B virus
c 183	15	1.0	17	21	AAA79915	Hepatitis B virus
184	15	1.0	18	18	AAX73480	Mouse flk-1 VEGF r
185	15	1.0	19	21	AAA83473	cdk8 ribozyme bind
186	15	1.0	19	21	AAA83474	cdk8 ribozyme bind
187	15	1.0	19	21	AAA83475	cdk8 ribozyme bind
188	15	1.0	19	21	AAA83476	cdk8 ribozyme bind
c 189	15	1.0	20	20	AAX96582	PCR primer used to
c 190	15	1.0	22	18	AAT51345	Primer 6 for alkal
191	15	1.0	24	18	AAT47831	Mouse NF2 gene PCR
192	15	1.0	24	20	AAX04297	Mouse NF2 transcri
c 193	15	1.0	30	19	AAV27211	Human alpha 1 coll
c 194	15	1.0	30	21	AAZ98716	Promoter region of
c 195	15	1.0	32	14	AAQ36232	Conserved polypuri
196	15	1.0	42	18	AAV13314	Primer BF25 for B.
c 197	15	1.0	43	14	AAQ36228	CAT box of human a
198	15	1.0	60	12	AAQ13356	Encodes ARH-III an
c 199	15	1.0	80	11	AAQ08622	HPV-16 capture oli
c 200	15	1.0	80	13	AAQ20226	Sequence of human
201	15	1.0	100	21	AAA63637	Luciferase gene ta
202	15	1.0	102	17	AAT30164	Nucleotides 7052-7
c 203	15	1.0	135	17	AAT07234	Immunogen DNA from
204	15	1.0	149	21	AAA63638	Luciferase gene ta
205	15	1.0	150	21	AAA63639	Luciferase gene ta
c 206	15	1.0	153	21	AAC04964	Human secreted pro
c 207	15	1.0	183	21	AAC16564	Human secreted pro
208	15	1.0	190	21	AAC18183	Human secreted pro
c 209	15	1.0	204	21	AAC07825	Human secreted pro
210	15	1.0	206	17	AAT09767	Tomato genomic DNA
211	15	1.0	214	16	AAT22996	Human gene signatu
212	15	1.0	217	21	AAC26128	Human secreted pro
213	15	1.0	243	21	AAA44176	Mouse secreted exp
214	15	1.0	250	21	AAA63640	Luciferase gene ta
c 215	15	1.0	251	21	AAC07741	Human secreted pro
216	15	1.0	252	16	AAT22216	Human gene signatu
c 217	15	1.0	261	20	AAZ12813	Human gene express
c 218	15	1.0	266	21	AAC18090	Human secreted pro
c 219	15	1.0	282	21	AAF10965	Fusarium venenatum
c 220	15	1.0	300	20	AAZ14245	Human gene express
221	15	1.0	300	20	AAZ13851	Human gene express
222	15	1.0	300	21	AAA00748	Human colon cancer
c 223	15	1.0	307	22	AAF93415	Spinal cord tissue
c 224	15	1.0	315	21	AAA45512	Human secreted exp
225	15	1.0	332	16	AAT26216	Human gene signatu
c 226	15	1.0	336	21	AAC01319	Human secreted pro
227	15	1.0	350	13	AAQ30076	Sequence contg. th
c 228	15	1.0	372	21	AAC94267	Cat flea head and
229	15	1.0	381	22	AAF64655	Novel human polynu
c 230	15	1.0	386	21	AAC00664	Human secreted pro

231	15	1.0	389	20	AAX41082	Human secreted pro
c 232	15	1.0	394	18	AAV77268	Staphylococcus aur
c 233	15	1.0	397	22	AAF65029	Novel human polynu
234	15	1.0	417	21	AAC75075	Human ORFX ORF630
c 235	15	1.0	424	21	AAC09482	Human secreted pro
236	15	1.0	428	20	AAV87491	EST clone BU60. H
c 237	15	1.0	428	21	AAC09483	Human secreted pro
c 238	15	1.0	436	14	AAQ45083	Human glucokinase
239	15	1.0	457	21	AAC01853	Human secreted pro
240	15	1.0	461	21	AAC53857	Arabidopsis thalia
c 241	15	1.0	462	20	AAV89989	EST clone CW691.
c 242	15	1.0	465	20	AAX37370	Human secreted pro
243	15	1.0	470	21	AAZ51882	N. tabacum Enh10 g
244	15	1.0	474	21	AAA68197	Pinus radiata pero
245	15	1.0	498	21	AAC93890	Cat flea hindgut a
c 246	15	1.0	498	21	AAC94510	Cat flea hindgut a
247	15	1.0	500	21	AAC94227	Cat flea head and
c 248	15	1.0	501	22	AAA54388	Plant defensin par
249	15	1.0	503	21	AAC19307	Human secreted pro
c 250	15	1.0	544	21	AAC00663	Human secreted pro
251	15	1.0	549	21	AAC93591	Cat flea head and
252	15	1.0	549	21	AAC94360	Cat flea head and
253	15	1.0	552	21	AAF16373	Human prostate can
c 254	15	1.0	554	21	AAC77796	Human cancer assoc
255	15	1.0	558	19	AAV19126	Nucleotide sequenc
c 256	15	1.0	562	21	AAA69837	Human ovarian carc
257	15	1.0	579	18	AAV01886	Human cytoskeleton
258	15	1.0	593	21	AAC36997	Arabidopsis thalia
259	15	1.0	595	18	AAT83987	DNA encoding a Sta
260	15	1.0	621	21	AAF09440	Fusarium venenatum
261	15	1.0	621	21	AAA65942	E. coli proliferat
262	15	1.0	632	21	AAC09063	Human secreted pro
c 263	15	1.0	636	21	AAZ94750	Human ATP binding
264	15	1.0	639	22	AAC84300	Human EXCS encodin
265	15	1.0	643	21	AAC53330	Arabidopsis thalia
c 266	15	1.0	658	22	AAA54389	Plant defensin cod
267	15	1.0	661	21	AAF14846	Aspergillus oryzae
268	15	1.0	662	21	AAF11078	Fusarium venenatum
269	15	1.0	675	21	AAC46118	Arabidopsis thalia
270	15	1.0	686	21	AAA02017	Human colon cancer
271	15	1.0	689	21	AAC34817	Arabidopsis thalia
272	15	1.0	689	21	AAC54064	Arabidopsis thalia
273	15	1.0	701	21	AAF13093	Aspergillus oryzae
274	15	1.0	703	21	AAF13188	Aspergillus oryzae
c 275	15	1.0	704	20	AAZ15948	Human gene express
c 276	15	1.0	720	21	AAZ97452	Human prostate can
c 277	15	1.0	724	20	AAX39663	Renal cancer assoc
c 278	15	1.0	738	21	AAZ97451	Human prostate can
279	15	1.0	742	21	AAA01753	Human colon cancer
c 280	15	1.0	747	21	AAA70085	Plasmodium falcipa
281	15	1.0	750	19	AAV52479	Streptococcus pneu
c 282	15	1.0	752	22	AAF22834	Human prostate can
283	15	1.0	754	20	AAZ16273	Human gene express
c 284	15	1.0	758	19	AAV18170	Gulf War veteran s
c 285	15	1.0	759	19	AAV18172	Gulf War veteran s
286	15	1.0	759	21	AAC40023	Arabidopsis thalia
c 287	15	1.0	760	19	AAV18171	Gulf War veteran s
288	15	1.0	767	16	AAQ85889	Wild-type bovine p
c 289	15	1.0	768	20	AAZ16930	Human gene express
290	15	1.0	770	21	AAA16537	Human colon cancer
c 291	15	1.0	777	22	AAF30008	Mouse CARD-5 cDNA.
c 292	15	1.0	793	22	AAF75346	Human TGF-beta rec
c 293	15	1.0	801	21	AAA70185	Plasmodium falcipa
294	15	1.0	807	20	AAZ42245	Human normal bladd
c 295	15	1.0	818	16	AAQ85098	PagC-pagD intergen
c 296	15	1.0	818	18	AAT70216	S. typhimurium pag
c 297	15	1.0	820	21	AAC54870	Arabidopsis thalia
c 298	15	1.0	824	21	AAC52453	Arabidopsis thalia
299	15	1.0	834	21	AAC48263	Arabidopsis thalia
300	15	1.0	836	21	AAC32695	Arabidopsis thalia
301	15	1.0	837	21	AAC54139	Arabidopsis thalia
c 302	15	1.0	840	18	AAV74819	Staphylococcus aur
c 303	15	1.0	841	21	AAF13198	Aspergillus oryzae
c 304	15	1.0	852	20	AAX39982	Prostate cancer as
305	15	1.0	852	21	AAC76629	Human ORFX ORF2184
306	15	1.0	862	20	AAX39858	Gastric cancer ass
c 307	15	1.0	882	21	AAC48988	Arabidopsis thalia
c 308	15	1.0	885	20	AAX39977	Prostate cancer as
c 309	15	1.0	887	20	AAZ07165	Human lung tumour
c 310	15	1.0	887	21	AAC79081	Human lung tumour-
c 311	15	1.0	902	16	AAQ81202	Cloned aldolase ge
c 312	15	1.0	916	21	AAC41805	Arabidopsis thalia
313	15	1.0	924	18	AAV06106	Viral infection ge
c 314	15	1.0	957	20	AAZ17643	Human gene express
c 315	15	1.0	968	21	AAC81100	Human secreted pro
316	15	1.0	972	20	AAZ17047	Human gene express
c 317	15	1.0	973	21	AAC33442	Arabidopsis thalia
318	15	1.0	978	21	AAC51996	Arabidopsis thalia
c 319	15	1.0	1001	21	AAC57458	Arachidonic acid m
c 320	15	1.0	1001	21	AAC57964	Arachidonic acid m
c 321	15	1.0	1001	21	AAC57965	Arachidonic acid m
c 322	15	1.0	1002	21	AAA09350	G protein-coupled
c 323	15	1.0	1012	21	AAC49486	Arabidopsis thalia
c 324	15	1.0	1014	21	AAC40326	Arabidopsis thalia
325	15	1.0	1015	18	AAV01887	Human cytoskeleton
326	15	1.0	1023	21	AAC45469	Arabidopsis thalia
c 327	15	1.0	1029	21	AAC34691	Arabidopsis thalia
328	15	1.0	1033	19	AAV60017	Nucleic acid E180c
329	15	1.0	1047	21	AAC33256	Arabidopsis thalia
330	15	1.0	1084	21	AAC54036	Arabidopsis thalia
331	15	1.0	1088	21	AAC52594	Arabidopsis thalia
c 332	15	1.0	1090	21	AAC40924	Arabidopsis thalia
333	15	1.0	1093	21	AAC51904	Arabidopsis thalia
334	15	1.0	1094	21	AAC36709	Arabidopsis thalia
c 335	15	1.0	1104	18	AAT84565	Swinepox virus thy
336	15	1.0	1105	21	AAC52035	Arabidopsis thalia
c 337	15	1.0	1109	3	AAN20026	Human leukocyte in
c 338	15	1.0	1110	21	AAC53744	Zea mays DNA fragm
339	15	1.0	1117	21	AAC42413	Arabidopsis thalia
c 340	15	1.0	1119	11	AAQ05129	Sequence encoding
c 341	15	1.0	1122	18	AAT62690	Lactate oxidase va
c 342	15	1.0	1122	18	AAT62691	Lactate oxidase va
c 343	15	1.0	1122	18	AAT62692	Lactate oxidase va
344	15	1.0	1131	21	AAZ53167	Neisseria meningit
345	15	1.0	1131	21	AAZ53168	Neisseria meningit
c 346	15	1.0	1132	21	AAC46901	Arabidopsis thalia
c 347	15	1.0	1150	19	AAV15112	SEQ ID NO:11 mRNA
c 348	15	1.0	1153	16	AAQ92410	Escherichia coli N
349	15	1.0	1155	19	AAV59893	DNA encoding Staph
350	15	1.0	1166	21	AAC39355	Arabidopsis thalia
c 351	15	1.0	1170	21	AAC38475	Arabidopsis thalia
352	15	1.0	1174	21	AAC49786	Arabidopsis thalia
c 353	15	1.0	1175	11	AAQ04088	Bo44-15 insert enc
354	15	1.0	1184	19	AAV19961	Synthetic lysyl ox
355	15	1.0	1192	21	AAC39980	Arabidopsis thalia
c 356	15	1.0	1194	21	AAZ53649	Neisseria gonorrhe
c 357	15	1.0	1194	21	AAZ53650	Neisseria meningit
c 358	15	1.0	1194	21	AAZ53651	Neisseria meningit
c 359	15	1.0	1199	18	AAT79691	BRCA2 cancer susce
c 360	15	1.0	1205	18	AAT92824	Flea saliva protei
c 361	15	1.0	1205	19	AAV73413	Flea saliva protei
362	15	1.0	1205	19	AAV73414	Flea saliva protei
c 363	15	1.0	1209	21	AAC40263	Arabidopsis thalia
c 364	15	1.0	1224	21	AAC35030	Arabidopsis thalia
365	15	1.0	1224	21	AAC52812	Arabidopsis thalia
366	15	1.0	1225	21	AAC54037	Arabidopsis thalia
c 367	15	1.0	1232	21	AAC73867	Human secreted pro
c 368	15	1.0	1235	14	AAQ33062	Encodes Babesia bo
c 369	15	1.0	1235	17	AAT18993	Babesia merozoite.
c 370	15	1.0	1240	20	AAX39976	Prostate cancer as
371	15	1.0	1259	21	AAC66428	Human secreted pro
c 372	15	1.0	1264	21	AAC73898	Human secreted pro
373	15	1.0	1275	21	AAZ90477	Cancer specific ge
374	15	1.0	1276	21	AAC77851	Human cancer assoc
375	15	1.0	1276	22	AAF58332	Human GTP-binding
376	15	1.0	1282	21	AAC48871	Arabidopsis thalia

c	377	15	1.0	1284	20	AAX20326	Borrelia burgdorfe
	378	15	1.0	1284	21	AAC37548	Arabidopsis thalia
c	379	15	1.0	1285	19	AAZ96387	S. pneumoniae deri
	380	15	1.0	1287	21	AAZ50357	Human CD39-L4 vari
	381	15	1.0	1287	22	AAF63385	DNA encoding human
	382	15	1.0	1287	22	AAF63402	Human CD39 like pr
	383	15	1.0	1302	21	AAZ94667	Human neuropeptide
c	384	15	1.0	1320	19	AAV44138	Chimeric rhPDGF-B
	385	15	1.0	1323	21	AAZ08254	Moraxella catarrha
	386	15	1.0	1323	21	AAZ08255	Moraxella catarrha
	387	15	1.0	1323	21	AAZ08256	Moraxella catarrha
	388	15	1.0	1323	21	AAZ08257	Moraxella catarrha
c	389	15	1.0	1332	21	AAC42359	Arabidopsis thalia
	390	15	1.0	1344	21	AAA80619	Human secreted pro
c	391	15	1.0	1344	21	AAC49477	Arabidopsis thalia
	392	15	1.0	1350	20	AAZ00364	Nucleotide sequenc
	393	15	1.0	1356	21	AAC51981	Arabidopsis thalia
c	394	15	1.0	1358	16	AAT00938	Almond N-glycosida
	395	15	1.0	1358	21	AAC37877	Arabidopsis thalia
c	396	15	1.0	1361	21	AAC44050	Zea mays DNA fragm
	397	15	1.0	1363	21	AAC40628	Arabidopsis thalia
	398	15	1.0	1370	21	AAC48945	Arabidopsis thalia
	399	15	1.0	1373	21	AAC42135	Arabidopsis thalia
	400	15	1.0	1377	21	AAA51791	S. cerevisiae esse
	401	15	1.0	1379	21	AAC33152	Arabidopsis thalia
	402	15	1.0	1380	17	AAT72048	Sequence around ma
	403	15	1.0	1380	17	AAT43928	Sequence around ma
c	404	15	1.0	1396	20	AAZ21259	Human cell junctio
	405	15	1.0	1399	16	AAQ89344	Human mGluR7 clone
	406	15	1.0	1403	21	AAC50494	Arabidopsis thalia
	407	15	1.0	1406	21	AAC39988	Arabidopsis thalia
c	408	15	1.0	1426	16	AAT00937	Almond N-glycosida
	409	15	1.0	1432	21	AAC42309	Arabidopsis thalia
c	410	15	1.0	1435	16	AAT00935	Almond N-glycosida
	411	15	1.0	1436	21	AAC79835	Human secreted pro
	412	15	1.0	1440	21	AAA57839	DNA encoding a hum
	413	15	1.0	1443	21	AAC35345	Arabidopsis thalia
	414	15	1.0	1446	21	AAA52649	Eosinophil activat
	415	15	1.0	1448	22	AAF94996	Human ovarian canc
	416	15	1.0	1474	17	AAT15759	Human interleukin-
	417	15	1.0	1484	20	AAX89755	Probe sequence for
c	418	15	1.0	1497	22	AAF60794	Psudomonas sp ABC
	419	15	1.0	1508	20	AAV55740	Partial human 7-tr
	420	15	1.0	1508	21	AAA61231	DNA encoding HLWAR
	421	15	1.0	1510	17	AAT15760	Human interleukin-
	422	15	1.0	1517	14	AAQ47166	BPV1 L1 ORF. Bovi
	423	15	1.0	1517	14	AAQ47167	HPV1 L1 ORF. Huma
	424	15	1.0	1517	19	AAV23923	HPV16 L1 gene #1.
	425	15	1.0	1517	19	AAV24099	HPV16 L1 gene #2.
	426	15	1.0	1517	19	AAV25817	Human papillomavir
	427	15	1.0	1517	19	AAV25818	Human papillomavir
	428	15	1.0	1517	19	AAV09855	HPV16 L1 capsid pr
	429	15	1.0	1517	19	AAV09856	HPV16 recombinant
	430	15	1.0	1517	19	AAV12162	Human papillomavir
	431	15	1.0	1517	19	AAV12161	Human papillomavir
	432	15	1.0	1517	20	AAX15111	Human papillomavir
	433	15	1.0	1517	20	AAX15112	Human papillomavir
	434	15	1.0	1517	21	AAZ40563	HPV16 L1 gene - co
	435	15	1.0	1517	21	AAZ40564	HPV16 L1 gene. Hu
	436	15	1.0	1518	20	AAX37566	Human papilloma vi
	437	15	1.0	1518	21	AAZ48174	Human papillomavir
	438	15	1.0	1521	12	AAQ14288	Human neuronal nic
	439	15	1.0	1534	17	AAT15995	Human interleukin-
c	440	15	1.0	1541	21	AAZ52488	Human secreted pro
c	441	15	1.0	1583	19	AAV32455	Human receptor tyr
	442	15	1.0	1601	21	AAZ50358	Human CD39-L66 pro
	443	15	1.0	1601	22	AAF63384	Human cDNA encodin
	444	15	1.0	1610	21	AAF21346	Human low adenosin
	445	15	1.0	1610	21	AAA35224	Human adenosine re
	446	15	1.0	1611	21	AAC77665	Human cancer assoc
	447	15	1.0	1641	17	AAT15758	Human interleukin-
c	448	15	1.0	1656	20	AAX25715	Firefly luciferase
c	449	15	1.0	1662	18	AAT88243	Modified firefly l
c	450	15	1.0	1665	18	AAT88244	Modified firefly l
	451	15	1.0	1669	20	AAZ41319	Human normal ovari
	452	15	1.0	1669	21	AAA92207	Petunia hybrida au
c	453	15	1.0	1705	21	AAF18105	Lung cancer associ
c	454	15	1.0	1722	16	AAT05621	Luciferase luc gen
c	455	15	1.0	1722	17	AAT36492	P. pyralis variant
c	456	15	1.0	1722	20	AAX02223	P. pyralis luc gen
c	457	15	1.0	1723	21	AAC98162	Human colon cancer
	458	15	1.0	1736	21	AAZ99988	DNA encoding the c
c	459	15	1.0	1747	13	AAQ20260	ERK1. Rat rattus.
c	460	15	1.0	1749	21	AAC43038	Arabidopsis thalia
	461	15	1.0	1752	22	AAF58315	Human GTP-binding
c	462	15	1.0	1767	21	AAC48989	Arabidopsis thalia
c	463	15	1.0	1768	21	AAC43652	Arabidopsis thalia
	464	15	1.0	1772	18	AAT62811	Cyclin-dependent k
	465	15	1.0	1776	18	AAT51339	Coding sequence fo
	466	15	1.0	1776	20	AAZ41331	Human normal uteru
	467	15	1.0	1786	21	AAC66234	Mutant alpha-amyla
c	468	15	1.0	1798	20	AAX06815	Firefly luciferase
	469	15	1.0	1799	21	AAZ50356	Human CD39-L4 prot
	470	15	1.0	1799	22	AAF63383	Human cDNA encodin
	471	15	1.0	1803	18	AAT87923	Rat cerebellum der
c	472	15	1.0	1811	15	AAQ58732	Encodes firefly lu
	473	15	1.0	1814	19	AAV82786	Clone er381_1 isol
c	474	15	1.0	1816	16	AAQ78896	VHL disease gene g
c	475	15	1.0	1816	18	AAT90895	Von Hippel-Lindau
	476	15	1.0	1820	21	AAA40755	Wild type human CD
	477	15	1.0	1825	19	AAV17601	Yeast AFC1 gene en
c	478	15	1.0	1844	13	AAQ30175	63 kD CaM PDE cDNA
c	479	15	1.0	1844	16	AAQ83969	Cyclic-GMP stimula
c	480	15	1.0	1844	18	AAT67200	Bovine brain 63 kD
c	481	15	1.0	1844	18	AAT51105	cDNA for 63 kD cal
c	482	15	1.0	1844	19	AAV48196	Bovine brain Ca2+/
c	483	15	1.0	1844	19	AAV54753	63 kDa CaM-PDE DNA
c	484	15	1.0	1844	19	AAV36148	Bovine CaM-PDE cDN
c	485	15	1.0	1844	21	AAZ90372	Bovine brain 63 kD
c	486	15	1.0	1850	22	AAC87116	Nucleotide sequenc
	487	15	1.0	1859	21	AAF16056	Human prostate can
	488	15	1.0	1859	21	AAD00013	Scorpion protein d
c	489	15	1.0	1862	18	AAT47814	Mouse NF2 cDNA iso
c	490	15	1.0	1862	20	AAX04294	Mouse neurofibroma
	491	15	1.0	1866	22	AAF29456	Human TANGO 325 cD
	492	15	1.0	1888	20	AAV68484	Human 7-transmembr
	493	15	1.0	1888	21	AAA61230	DNA encoding HLWAR
	494	15	1.0	1903	21	AAZ86955	Human NPY-Y7 recep
	495	15	1.0	1907	19	AAV28462	Synthetic gene enc
	496	15	1.0	1907	19	AAV29128	DNA sequence of a
	497	15	1.0	1917	21	AAA37848	Human obesity prot
c	498	15	1.0	1939	12	AAQ10026	Encodes a portion
	499	15	1.0	1961	18	AAT97973	DNA encoding canin
	500	15	1.0	1964	20	AAX00640	Human secreted pro
c	501	15	1.0	1965	13	AAQ25892	Sequence encodes e
	502	15	1.0	1998	21	AAA96068	Human ecto-phospha
	503	15	1.0	2032	17	AAT10117	Drosophila scaveng
	504	15	1.0	2056	22	AAF32785	Human secreted pro
c	505	15	1.0	2100	21	AAC48917	Arabidopsis thalia
c	506	15	1.0	2115	21	AAC32824	Arabidopsis thalia
c	507	15	1.0	2127	22	AAF33075	Human secreted pro
	508	15	1.0	2136	22	AAC82671	Human 7TM clone HD
c	509	15	1.0	2144	21	AAC39497	Arabidopsis thalia
	510	15	1.0	2149	21	AAC44178	Arabidopsis thalia
	511	15	1.0	2152	21	AAA59320	cDNA encoding a ca
c	512	15	1.0	2157	21	AAC54444	Arabidopsis thalia
c	513	15	1.0	2160	21	AAC52971	Arabidopsis thalia
	514	15	1.0	2169	22	AAF29455	Human TANGO 325 cD
c	515	15	1.0	2174	20	AAX01282	Human DSCR1 coding
	516	15	1.0	2205	16	AAQ97799	Clone pTzeta12 cod
	517	15	1.0	2240	21	AAA95062	cDNA from the rice
	518	15	1.0	2290	19	AAV43879	Mycobacterial meth
	519	15	1.0	2295	21	AAA80610	Human CD33-like se
	520	15	1.0	2320	13	AAQ26694	pag C. Salmonella
	521	15	1.0	2320	13	AAQ27739	pagC/AP fusion reg
	522	15	1.0	2320	16	AAQ85099	PagC gene. Salmon

523	15	1.0	2320	18	AAT70212	S. typhimurium pag
c 524	15	1.0	2334	11	AAQ06554	92-kDa human type
c 525	15	1.0	2334	19	AAV04528	Human gelatinase B
c 526	15	1.0	2363	20	AAZ27430	Bovine rotavirus I
527	15	1.0	2371	21	AAF21308	Human low adenosin
528	15	1.0	2371	21	AAA35186	Human adenosine re.
c 529	15	1.0	2382	21	AAC50567	Arabidopsis thalia
530	15	1.0	2391	17	AAT29215	Cephalosporin amid
531	15	1.0	2400	14	AAQ49353	deg-3. Panagrellu
532	15	1.0	2431	20	AAX58739	DNA encoding anti-
c 533	15	1.0	2447	13	AAQ30000	Chimeric lux::npt
534	15	1.0	2448	15	AAV12201	Human neuronal nic
535	15	1.0	2448	18	AAT48240	Neuronal nicotinic
c 536	15	1.0	2459	21	AAC75871	Human ORFX ORF1426
c 537	15	1.0	2465	14	AAQ51085	H. somnus lppC gen
c 538	15	1.0	2478	10	AAN91246	Sequence of B. thu
c 539	15	1.0	2481	21	AAC48385	Arabidopsis thalia
540	15	1.0	2482	21	AAC48600	Arabidopsis thalia
541	15	1.0	2528	21	AAA08039	Human protein tran
c 542	15	1.0	2534	21	AAC69141	Human ABC1 gene ex
c 543	15	1.0	2540	14	AAQ37687	Rat choline kinase
c 544	15	1.0	2580	21	AAC42792	Arabidopsis thalia
545	15	1.0	2584	21	AAC79615	Virulence gene #35
c 546	15	1.0	2619	21	AAZ52571	Human secreted pro
c 547	15	1.0	2707	20	AAX13010	Enterococcus faeca
c 548	15	1.0	2733	10	AAN91700	Human type IV coll
c 549	15	1.0	2750	14	AAQ38808	G6PD coding sequen
550	15	1.0	2764	20	AAZ34109	Human PRO940 nucle
551	15	1.0	2764	21	AAC78510	Human PRO940 (UNQ4
c 552	15	1.0	2781	21	AAA87991	Kapok CPE6-1A 5' u
c 553	15	1.0	2811	21	AAC76229	Human ORFX ORF1784
554	15	1.0	2816	12	AAQ14380	Soluble Insulin-li
555	15	1.0	2816	12	AAQ14814	Soluble human IGF-
556	15	1.0	2820	12	AAQ14816	Hybrid human insul
c 557	15	1.0	2826	17	AAT38071	Transferrin bindin
c 558	15	1.0	2850	21	AAC76428	Human ORFX ORF1983
c 559	15	1.0	2919	15	AAQ73750	Rice starch branch
560	15	1.0	2929	21	AAA46558	cDNA sequence enco
c 561	15	1.0	2966	20	AAX33415	Human Ras-like pro
c 562	15	1.0	2983	21	AAZ60809	Nucleotide sequenc
563	15	1.0	2995	20	AAV64417	Mouse developing l
c 564	15	1.0	3072	20	AAX28142	CCoAOMT promoter,
565	15	1.0	3076	19	AAV43674	Receptor type tyro
566	15	1.0	3079	20	AAZ23025	Rat kd312 genomic
567	15	1.0	3099	20	AAV99912	Human sialoadhesin
568	15	1.0	3099	20	AAV99911	Human sialoadhesin
c 569	15	1.0	3115	21	AAC98113	Human colon cancer
570	15	1.0	3165	18	AAT84001	DNA encoding a Sta
571	15	1.0	3165	19	AAV53408	DNA encoding 2 Sta
572	15	1.0	3253	21	AAZ29719	T. cruzi complemen
573	15	1.0	3261	22	AAF26571	DNA encoding human
c 574	15	1.0	3274	20	AAX13539	Enterococcus faeca
c 575	15	1.0	3306	11	AAQ05620	Expressible sequen
576	15	1.0	3312	8	AAN71034	pMx34 cDNA insert.
577	15	1.0	3329	22	AAD02458	Mouse C3a receptor
578	15	1.0	3347	22	AAF68400	Human lung tumour
579	15	1.0	3441	18	AAT87922	Rat cerebellum der
580	15	1.0	3505	12	AAQ11850	Glutamate receptor
581	15	1.0	3619	15	AAQ44685	Excitatory amino a
c 582	15	1.0	3625	21	AAF21034	Human low adenosin
c 583	15	1.0	3625	21	AAA34912	Human adenosine re
584	15	1.0	3627	15	AAQ70824	Protein kinase (NU
585	15	1.0	3627	16	AAQ92959	Saccharomyces cere
586	15	1.0	3628	20	AAV71905	S. cerevisiae CKI
587	15	1.0	3642	6	AAN50364	DNA fragment encod
588	15	1.0	3678	15	AAQ53996	Equine herpes viru
589	15	1.0	3678	20	AAX38299	Equine herpes viru
c 590	15	1.0	3691	20	AAZ06408	Tumour suppressor
c 591	15	1.0	3691	20	AAX99691	Nucleic acid seque
592	15	1.0	3738	20	AAX13645	Enterococcus faeca
c 593	15	1.0	3754	20	AAX59767	SEQ ID 6 of WO9919
594	15	1.0	3804	16	AAQ89343	Human mGluR7 clone
595	15	1.0	3825	20	AAX91484	Human CIF130 prote
596	15	1.0	3825	21	AAZ29604	Human cell cycle r
c 597	15	1.0	3869	21	AAA57993	3868 bp Candida al
c 598	15	1.0	3937	18	AAT75036	Chick fringe A (ra
599	15	1.0	3954	21	AAA75647	Nucleotide sequenc
600	15	1.0	4180	20	AAZ32010	Human METH1 relate
601	15	1.0	4180	22	AAC90067	D67076 cDNA clone.
602	15	1.0	4257	20	AAX98030	Nucleic acid encod
603	15	1.0	4371	20	AAX06877	Caenorhabditis ele
c 604	15	1.0	4372	18	AAT91360	Orf virus genomic
c 605	15	1.0	4500	13	AAQ29684	CCV-6 spike gene.
c 606	15	1.0	4505	21	AAZ45145	Matrix metalloprot
c 607	15	1.0	4656	22	AAF31895	Candida albicans A
608	15	1.0	4673	13	AAQ27189	P. yoelii SSP2 ant
609	15	1.0	4696	17	AAT13330	Rat type I insulin
610	15	1.0	4811	21	AAD00554	Human APM1 (Adipos
c 611	15	1.0	4846	20	AAV72416	P. pyralis pTetLuc
612	15	1.0	4864	21	AAZ94737	Human ATP binding
613	15	1.0	4868	19	AAV26068	Human canalicular
614	15	1.0	4975	15	AAQ73445	IGF-I receptor. H
c 615	15	1.0	4987	15	AAQ65322	Vaccinia virus pro
c 616	15	1.0	4987	16	AAQ86799	Vaccinia virus vec
c 617	15	1.0	4987	21	AAA29136	pNF-kappaB-Luc con
618	15	1.0	4989	15	AAQ73702	Human IFG-1 recept
619	15	1.0	4989	17	AAQ99245	Insulin like growt
620	15	1.0	4989	17	AAT29608	IGF-1 receptor gen
621	15	1.0	4989	17	AAT13329	Human type I insul
622	15	1.0	4989	18	AAT95673	Human insulin-like
623	15	1.0	4989	19	AAV27035	Homo sapiens IGF-1
624	15	1.0	4989	22	AAF45144	IGF-1 receptor cod
c 625	15	1.0	5049	19	AAV42322	Helicobacter pylor
626	15	1.0	5051	11	AAQ05682	Islets of Langerha
627	15	1.0	5051	20	AAZ32334	Human pancreatic i
628	15	1.0	5051	20	AAZ09915	Human islet cell a
629	15	1.0	5051	20	AAV63553	Islet cell antibod
630	15	1.0	5051	21	AAZ46452	Pancreatic islet c
c 631	15	1.0	5051	22	AAC86954	Nucleotide sequenc
c 632	15	1.0	5099	19	AAV38239	Nucleotide sequenc
c 633	15	1.0	5145	22	AAF24991	Nucleotide sequenc
c 634	15	1.0	5256	21	AAA07776	DNA sequence of pl
635	15	1.0	5404	17	AAT38735	Murine foetal live
636	15	1.0	5404	20	AAX77516	Murine flk-1 cDNA.
c 637	15	1.0	5405	22	AAC83389	Luciferase reporte
638	15	1.0	5406	13	AAQ29957	flk-1 cDNA sequenc
639	15	1.0	5406	14	AAQ53504	Murine flk-1 cDNA.
640	15	1.0	5406	14	AAQ35251	Human flk-1 coding
641	15	1.0	5406	14	AAQ40916	Murine flk-1 cDNA.
642	15	1.0	5406	16	AAQ81014	Flk1 receptor prot
643	15	1.0	5406	16	AAQ79070	Mouse flk-1 cDNA.
644	15	1.0	5406	18	AAT72119	Murine flk-1 recep
c 645	15	1.0	5427	22	AAC83390	Luciferase reporte
646	15	1.0	5470	15	AAQ64049	Sequence of murine
c 647	15	1.0	5559	15	AAQ55260	Restriction fragme
c 648	15	1.0	5559	18	AAT99212	5.6 kb E11 Pseudom
c 649	15	1.0	5559	18	AAT96101	Pseudomonas fluore
c 650	15	1.0	5559	18	AAT90945	E11 genomic DNA fr
c 651	15	1.0	5559	19	AAV06486	5.6 kb EcoRI-Hind
c 652	15	1.0	5559	20	AAX02982	Human IL-1ra BAC c
653	15	1.0	5586	18	AAT94023	Human canalicular
c 654	15	1.0	5595	13	AAQ25810	Insert of pSRQ220
c 655	15	1.0	5620	18	AAT48630	Ad5-ITR, CMV-lucif
c 656	15	1.0	5620	22	AAF30233	Minimal adenovirus
657	15	1.0	5648	21	AAZ46134	Endogenous plasmid
c 658	15	1.0	5696	20	AAX37304	Human hTLP2 cDNA.
c 659	15	1.0	5789	20	AAX08779	pGL2-Promoter gene
c 660	15	1.0	5791	20	AAX08778	pGL2(GST-ARE)SV40/
c 661	15	1.0	5793	20	AAX08776	pGL2(apo AI-ARE)SV
c 662	15	1.0	5793	20	AAX08777	pGL2(apo AI-mARE)S
c 663	15	1.0	5818	18	AAT85712	Human apolipoprote
c 664	15	1.0	5818	21	AAZ38352	Plasmid pGL2 (apoA
c 665	15	1.0	5819	18	AAT85711	Human apolipoprote
c 666	15	1.0	5819	20	AAX08775	pGL2(apo AI-491 mu
c 667	15	1.0	5819	20	AAX08774	pGL2(apo AI-250)lu
c 668	15	1.0	5819	21	AAZ38351	Plasmid pGL2 (apoA

c 669	15	1.0	5849	19	AAV33135	Plasmodium berghei
c 670	15	1.0	5874	20	AAX59765	cDNA encoding huma
c 671	15	1.0	5934	19	AAV68385	Human BAZ1-alpha c
c 672	15	1.0	5938	18	AAT85713	Human apolipoprote
c 673	15	1.0	5938	21	AAZ38353	Plasmid pGL2 (4xap
c 674	15	1.0	5960	19	AAV45225	Human nerve growth
675	15	1.0	6004	21	AAA07835	Human homologue of
c 676	15	1.0	6044	17	AAT28280	Plasmid pMV361-lux
677	15	1.0	6047	17	AAT28278	Plasmid pMH30-lux
c 678	15	1.0	6070	21	AAZ98335	A. thaliana gene i
c 679	15	1.0	6092	18	AAT85710	Human apolipoprote
c 680	15	1.0	6092	20	AAX08773	pGL2(apo AI-491)lu
c 681	15	1.0	6092	21	AAZ38350	Plasmid pGL2 (apoA
682	15	1.0	6121	21	AAA07779	DNA sequence of pl
c 683	15	1.0	6170	16	AAQ98929	Vector p19LUC sens
c 684	15	1.0	6171	17	AAT28279	Plasmid pMV261-lux
c 685	15	1.0	6229	16	AAQ98931	Vector p39LUC sens
686	15	1.0	6236	19	AAX75759	Human neurofilamen
c 687	15	1.0	6245	16	AAQ98915	TGF-beta responsiv
c 688	15	1.0	6254	16	AAQ98916	TGF-beta responsiv
c 689	15	1.0	6254	16	AAQ98918	TGF-beta responsiv
c 690	15	1.0	6265	16	AAQ98917	TGF-beta responsiv
c 691	15	1.0	6272	18	AAV74451	Staphylococcus aur
c 692	15	1.0	6464	16	AAQ98933	Vector HBVLuc. Sy
c 693	15	1.0	6565	22	AAF55126	Nucleotide sequenc
694	15	1.0	6595	18	AAT94659	Arabidopsis flavon
695	15	1.0	6831	18	AAT80200	Phosphatidyl inosi
c 696	15	1.0	6971	22	AAF55124	Nucleotide sequenc
697	15	1.0	7070	16	AAQ86897	S. cerevisiae fks2
c 698	15	1.0	7179	20	AAZ21006	Salmonella dublin
c 699	15	1.0	7223	18	AAT62602	Luciferase express
700	15	1.0	7294	20	AAX13000	Enterococcus faeca
701	15	1.0	7368	11	AAQ04013	Acetyl-CoA-carboxy
c 702	15	1.0	7558	22	AAF55125	Nucleotide sequenc
c 703	15	1.0	7577	19	AAV52179	Streptococcus pneu
704	15	1.0	7655	21	AAA07847	Hs-UNC-53/1 fragm
705	15	1.0	7900	20	AAX13068	Enterococcus faeca
706	15	1.0	7902	17	AAT09847	Human papilloma vi
707	15	1.0	7904	18	AAT94723	Human papillomavir
708	15	1.0	7904	18	AAT94724	Human papillomavir
709	15	1.0	7904	20	AAX33881	HPV-16 genomic seq
710	15	1.0	7904	20	AAV99946	Nucleotide sequenc
711	15	1.0	7904	20	AAV99952	Nucleotide sequenc
c 712	15	1.0	7969	22	AAF55123	Nucleotide sequenc
c 713	15	1.0	8220	16	AAQ83529	P. falciparum Proj
c 714	15	1.0	8220	18	AAT72897	Plasmodium Proj3 g
c 715	15	1.0	8220	21	AAZ98286	P. falciparum Proj
716	15	1.0	9248	20	AAZ32011	Human METH1 relate
717	15	1.0	9248	22	AAC90068	AB001735 cDNA clon
718	15	1.0	9365	21	AAZ50359	Human CD39-L4 geno
719	15	1.0	9365	22	AAF63405	Human CD39 like pr
720	15	1.0	9388	18	AAX83330	Genomic sequence o
721	15	1.0	9388	18	AAT75001	Retroviral element
722	15	1.0	9388	19	AAV69044	Human endogenous r
723	15	1.0	9388	21	AAC80853	Human B18Ag1 genom
724	15	1.0	9717	20	AAZ06826	Photorhabdus lumin
c 725	15	1.0	10160	20	AAV63723	Vector plasmid pDA
726	15	1.0	10329	21	AAZ50264	Genomic DNA of pot
727	15	1.0	10329	21	AAZ50653	Genomic DNA encodi
c 728	15	1.0	10533	16	AAQ98930	Vector p39neoLUC s
c 729	15	1.0	10549	16	AAQ98911	TGF-beta responsiv
c 730	15	1.0	10558	16	AAQ98912	TGF-beta responsiv
c 731	15	1.0	10558	16	AAQ98914	TGF-beta responsiv
c 732	15	1.0	10569	16	AAQ98913	TGF-beta responsiv
c 733	15	1.0	10697	16	AAQ98910	TGF-beta responsiv
c 734	15	1.0	10708	19	AAV69286	Sequence of mouse
735	15	1.0	10754	20	AAZ28236	Human adenovirus r
736	15	1.0	10754	20	AAX86142	Partial sequence o
c 737	15	1.0	10763	19	AAV39837	Pseudomonas fluore
c 738	15	1.0	10763	20	AAX99368	P. fluorescens 11k
c 739	15	1.0	10768	16	AAQ98932	Vector HBVneoLuc.
740	15	1.0	11266	20	AAX12991	Enterococcus faeca
c 741	15	1.0	11293	16	AAQ98909	TGF-beta responsiv
c 742	15	1.0	11459	20	AAX06876	Caenorhabditis ele
c 743	15	1.0	11462	19	AAV05532	Nucleic acid seque
c 744	15	1.0	11473	20	AAX34650	Starch branching e
c 745	15	1.0	11601	12	AAQ13608	ACV synthetase gen
c 746	15	1.0	11616	14	AAQ38105	Vector pHEBO-DR-LU
c 747	15	1.0	11616	15	AAQ67390	Construct pHEBO-DR
748	15	1.0	11721	20	AAZ07666	A. thaliana APETAL
c 749	15	1.0	11721	20	AAZ07666	A. thaliana APETAL
750	15	1.0	11781	21	AAA51392	Chromosome 16q tum
c 751	15	1.0	11784	20	AAV63724	Vector plasmid pDA
c 752	15	1.0	11785	15	AAQ74694	Genomic RNA sequen
c 753	15	1.0	11785	20	AAX03390	Entire sequence of
c 754	15	1.0	11970	20	AAX75923	Human interleukin
c 755	15	1.0	11970	21	AAF20943	Human interleukin-
c 756	15	1.0	11970	21	AAC63767	Human IL-1A gene.
c 757	15	1.0	11970	21	AAA34821	Human adenosine re
c 758	15	1.0	11970	22	AAF27665	IL-1A DNA. Uniden
c 759	15	1.0	11970	22	AAC91433	Human IL-1A nucleo
c 760	15	1.0	11991	20	AAV63725	Vector plasmid pDA
761	15	1.0	12117	21	AAA96368	Polymorphic repeat
c 762	15	1.0	12614	22	AAC66931	Plant signal trans
c 763	15	1.0	12886	21	AAA09084	Human patched-like
c 764	15	1.0	13058	14	AAQ48231	Vector containing
765	15	1.0	13417	17	AAT09224	Partial sequence o
c 766	15	1.0	13425	19	AAV52284	Streptococcus pneu
c 767	15	1.0	13548	21	AAF21037	Human low adenosin
c 768	15	1.0	13548	21	AAA34915	Human adenosine re
c 769	15	1.0	13574	21	AAA81529	N. meningitidis pa
c 770	15	1.0	13940	22	AAF25915	Human tumor suppre
c 771	15	1.0	14113	21	AAA90096	PR-1 promoter/luci
c 772	15	1.0	14113	21	AAA39651	Fusion construct P
773	15	1.0	14176	18	AAT84564	Swinepox virus Hin
c 774	15	1.0	14194	22	AAC66932	Plant signal trans
775	15	1.0	14636	22	AAF31864	Human DNA-PKcs cod
776	15	1.0	14698	19	AAV61792	Marker plasmid Des
777	15	1.0	14747	22	AAF63406	Human CD39 like pr
c 778	15	1.0	15186	21	AAZ44617	Newcastle disease
779	15	1.0	15977	22	AAF63407	Human CD39 like pr
c 780	15	1.0	17041	13	AAQ21065	Genomic DNA of hum
c 781	15	1.0	17041	15	AAQ67057	PACPAP38 DNA. Hom
782	15	1.0	17082	20	AAX13166	Enterococcus faeca
783	15	1.0	19001	19	AAV61793	Traget plasmid Mol
784	15	1.0	19035	19	AAV61794	Traget plasmid Man
c 785	15	1.0	19124	18	AAT72882	Plasmodium var-7 g
c 786	15	1.0	19124	21	AAZ98287	Plasmodium var-7 p
787	15	1.0	20633	20	AAX13213	Enterococcus faeca
788	15	1.0	20966	21	AAD00552	Human APM1 (Adipos
c 789	15	1.0	24978	20	AAX60209	SEQ ID 3 of US5914
c 790	15	1.0	24979	21	AAA52321	Genomic DNA sequen
c 791	15	1.0	28866	20	AAX22304	Human IL-1ra BAC c
c 792	15	1.0	29392	19	AAV15422	Mouse poly Ig rece
c 793	15	1.0	29433	21	AAF20950	Human interleukin-
c 794	15	1.0	29433	21	AAA34828	Human adenosine re
795	15	1.0	30246	18	AAV74367	Staphylococcus aur
c 796	15	1.0	32042	20	AAZ09252	Human CARD-4 DNA.
c 797	15	1.0	32042	22	AAF30011	Human CARD-4 gene.
798	15	1.0	36519	19	AAV22141	Chimpanzee adenovi
799	15	1.0	40298	21	AAF21311	Human low adenosin
800	15	1.0	40298	21	AAA35189	Human adenosine re
801	15	1.0	41708	21	AAA68247	Bacteriophage 77 c
802	15	1.0	43676	21	AAZ60606	Nucleotide sequenc
803	15	1.0	43795	21	AAZ92583	Human DAZ genomic
c 804	15	1.0	46899	15	AAQ54386	T. niveum Cyclospo
c 805	15	1.0	49646	21	AAA81457	N. meningitidis pa
c 806	15	1.0	49999	20	AAZ23901	Human LOBO homolog
807	15	1.0	51952	19	AAV26084	Tomato pest resist
c 808	15	1.0	52253	21	AAA81478	N. meningitidis pa
c 809	15	1.0	53526	19	AAT94101	Human PKD1 gene.
c 810	15	1.0	53577	17	AAT18551	Human polycystic k
c 811	15	1.0	53577	19	AAT94108	Human PKD1 locus b
812	15	1.0	59590	21	AAF22281	BAC containing rep
813	15	1.0	60873	21	AAA81469	N. meningitidis pa
c 814	15	1.0	60873	21	AAA81469	N. meningitidis pa

c 815	15	1.0	69936	21	AAA81479	N. meningitidis pa
c 816	15	1.0	72604	20	AAZ10752	Genomic sequence o
817	15	1.0	81145	22	AAF54868	Genomic nucleotide
c 818	15	1.0	84607	20	AAX90847	Human PACAP genomi
c 819	15	1.0	87350	18	AAX83003	Human WRN genomic
820	15	1.0	94750	22	AAF28551	Genomic fragment #
c 821	15	1.0	96109	22	AAF28548	Genomic fragment #
822	15	1.0	99960	21	AAZ50905	Human TBC-1 partia
823	15	1.0	100848	22	AAF28552	Genomic fragment #
824	15	1.0	106746	21	AAA10225	Human PCTA-1 genom
c 825	15	1.0	106746	21	AAA10225	Human PCTA-1 genom
c 826	15	1.0	116277	20	AAX20249	Borrelia burgdorfe
827	15	1.0	151826	21	AAF22291	BAC containing rep
c 828	15	1.0	151826	21	AAF22291	BAC containing rep
c 829	15	1.0	160552	22	AAD02697	Human glycosyl sul
830	15	1.0	162450	21	AAZ86967	Retinoblastoma bin
831	15	1.0	172325	21	AAF21613	Neisseria meningit
c 832	15	1.0	183999	22	AAF92831	Human ABC1 genomic
c 833	15	1.0	209273	21	AAF21437	Human factor-relat
c 834	15	1.0	269223	22	AAF28554	Genomic fragment #
835	15	1.0	349980	21	AAF21544	Neisseria meningit
836	15	1.0	349980	21	AAF21608	Neisseria meningit
837	15	1.0	349980	21	AAF21610	Neisseria meningit
c 838	15	1.0	349980	21	AAF21610	Neisseria meningit
839	15	1.0	580073	18	AAT58840	Mycoplasma genital
c 840	15	1.0	611590	21	AAF22303	Arabidopsis thalia
841	15	1.0	910715	20	AAX20248	Borrelia burgdorfe
842	15	1.0	1038602	20	AAZ01425	Complete genome se
c 843	15	1.0	1038602	20	AAZ01425	Complete genome se
844	15	1.0	1082138	21	AAF22305	Arabidopsis thalia
845	15	1.0	1230025	20	AAX91990	Nucleotide sequenc
c 846	15	1.0	1230025	20	AAX91990	Nucleotide sequenc
c 847	15	1.0	1664976	19	AAV21209	Methanococcus jann
c 848	14	0.9	15	16	AAT54205	Human IL-5 hammerh
849	14	0.9	15	22	AAF51815	IGF-I oligonucleot
850	14	0.9	15	22	AAF51817	IGF-I oligonucleot
c 851	14	0.9	17	16	AAQ94360	Septoria nodorum I
852	14	0.9	17	19	AAV96654	Potato citrate syn
853	14	0.9	17	19	AAV96655	Potato citrate syn
c 854	14	0.9	17	19	AAV62508	Septoria nodorum s
855	14	0.9	17	20	AAA18801	Human TIE-2 substr
856	14	0.9	18	19	AAV57104	Human Notch3 mutan
857	14	0.9	19	21	AAA83472	cdk8 ribozyme bind
c 858	14	0.9	19	21	AAA79913	Hepatitis B virus
c 859	14	0.9	20	21	AAA95322	B. cereus zwitterm
c 860	14	0.9	20	21	AAA74166	Reverse PCR primer
861	14	0.9	21	15	AAQ70747	3' primer for ampl
c 862	14	0.9	21	20	AAX24581	Human SR-BI gene i
863	14	0.9	21	20	AAX24583	Human SR-BI gene i
c 864	14	0.9	21	20	AAX24673	Human SR-BI gene i
865	14	0.9	21	20	AAX24675	Human SR-BI gene i
c 866	14	0.9	23	16	AAQ94361	Septoria nodorum I
c 867	14	0.9	23	19	AAV62509	Septoria nodorum s
c 868	14	0.9	25	21	AAC95628	HLA DQB gene PCR p
c 869	14	0.9	25	21	AAC95640	HLA DQB gene PCR p
c 870	14	0.9	25	21	AAC71006	Single nucleotide
c 871	14	0.9	28	18	AAV14396	Primer T221 for ca
c 872	14	0.9	29	18	AAV14397	Primer T222 for ca
c 873	14	0.9	29	21	AAZ49117	PCR primer for FIL
874	14	0.9	30	17	AAT33971	Inosine-guanosine
875	14	0.9	30	21	AAZ98644	Probe #2 used in h
876	14	0.9	30	21	AAZ98668	Probe #2 used in a
c 877	14	0.9	31	20	AAX24582	Human SR-BI gene i
878	14	0.9	31	20	AAX24584	Human SR-BI gene i
879	14	0.9	31	20	AAX24676	Human SR-BI gene i
c 880	14	0.9	31	20	AAX24674	Human SR-BI gene i
881	14	0.9	31	21	AAZ24470	N. meningitidis si
882	14	0.9	33	18	AAT85947	5' primer for ampl
c 883	14	0.9	36	16	AAT57605	RSV N hammerhead r
884	14	0.9	38	21	AAZ54951	Neisseria species
885	14	0.9	40	21	AAA76193	Escherichia coli i
886	14	0.9	42	11	AAQ05248	Probe to gene enco
887	14	0.9	42	19	AAV37922	Platelet derived e
888	14	0.9	50	15	AAQ69411	Human androgen rec
c 889	14	0.9	50	15	AAQ70977	GST2 gene specific
890	14	0.9	50	18	AAT63873	Human androgen rec
891	14	0.9	50	20	AAX17161	Test sequence from
c 892	14	0.9	51	21	AAA77056	Human clone cg4391
c 893	14	0.9	51	21	AAA77057	Human clone cg4391
c 894	14	0.9	53	19	AAV55833	Downstream frame s
c 895	14	0.9	56	18	AAV79348	Staphylococcus aur
c 896	14	0.9	57	21	AAC68884	Human interleukin-
897	14	0.9	58	22	AAF75670	Human beta-globin
898	14	0.9	59	21	AAC29956	Human secreted pro
c 899	14	0.9	60	20	AAX15726	PCR primer used to
900	14	0.9	66	13	AAQ33742	Upstream sequence
901	14	0.9	68	20	AAZ09553	Chimeric mutationa
902	14	0.9	68	20	AAZ09554	Chimeric mutationa
903	14	0.9	68	20	AAV72617	Chimeric mutationa
904	14	0.9	68	20	AAV72618	Chimeric mutationa
c 905	14	0.9	72	19	AAT97279	Nucleotide EP-16 S
906	14	0.9	72	19	AAT97280	Nucleotide EP-17 S
907	14	0.9	84	21	AAC13671	Human secreted pro
c 908	14	0.9	86	19	AAV41231	Mouse embryonic ce
c 909	14	0.9	87	16	AAT06064	Insulin receptor l
c 910	14	0.9	87	16	AAT06069	Insulin receptor l
911	14	0.9	89	20	AAX54641	Human IL-5 antisen
912	14	0.9	89	21	AAF20210	Human IL5 polynucl
913	14	0.9	89	21	AAA34088	Human adenosine re
914	14	0.9	96	20	AAX03315	Test RNA fragment
915	14	0.9	103	18	AAV77014	Staphylococcus aur
c 916	14	0.9	106	19	AAV09674	Human cathepsin K
c 917	14	0.9	107	18	AAV79143	Staphylococcus aur
918	14	0.9	108	21	AAZ99032	RNA aptamer #9 for
919	14	0.9	113	16	AAT23775	Human gene signatu
c 920	14	0.9	116	18	AAV77016	Staphylococcus aur
921	14	0.9	117	21	AAC32462	Human secreted pro
c 922	14	0.9	119	21	AAC20583	Human secreted pro
c 923	14	0.9	125	19	AAX10687	Human biallelic po
924	14	0.9	127	21	AAC26239	Human secreted pro
c 925	14	0.9	130	21	AAA45115	Human secreted exp
926	14	0.9	131	18	AAV00434	3' fragment of clo
c 927	14	0.9	131	21	AAC28695	Human secreted pro
c 928	14	0.9	136	21	AAC29577	Human secreted pro
c 929	14	0.9	143	21	AAC24132	Human secreted pro
c 930	14	0.9	145	17	AAT06755	Synthetic DNA enco
931	14	0.9	145	18	AAV77518	Staphylococcus aur
c 932	14	0.9	166	16	AAT21452	Human gene signatu
c 933	14	0.9	167	15	AAQ72535	Osteoclast-specifi
c 934	14	0.9	167	19	AAV04515	Osteoclast specifi
935	14	0.9	168	21	AAC19784	Human secreted pro
936	14	0.9	169	21	AAZ45692	DNA encoding a hel
c 937	14	0.9	175	21	AAC30235	Human secreted pro
938	14	0.9	176	21	AAC56221	Pinus radiata tran
939	14	0.9	177	21	AAC25196	Human secreted pro
c 940	14	0.9	178	21	AAF10719	Fusarium venenatum
941	14	0.9	181	13	AAQ24626	GSE C. Homo sapie
c 942	14	0.9	182	21	AAC20199	Human secreted pro
c 943	14	0.9	186	21	AAC25784	Human secreted pro
c 944	14	0.9	188	21	AAC31331	Human secreted pro
945	14	0.9	189	21	AAC16046	Human secreted pro
c 946	14	0.9	192	21	AAC04433	Human secreted pro
c 947	14	0.9	196	21	AAA45399	Human secreted exp
c 948	14	0.9	199	21	AAC05417	Human secreted pro
949	14	0.9	201	21	AAC26227	Human secreted pro
c 950	14	0.9	202	21	AAA45225	Human secreted exp
c 951	14	0.9	205	21	AAA44746	Human secreted exp
952	14	0.9	206	13	AAQ24630	GSE G. Homo sapie
c 953	14	0.9	209	20	AAX05191	Human MSH5 (hMSH5)
c 954	14	0.9	213	18	AAV78942	Staphylococcus aur
955	14	0.9	215	21	AAC71671	Single nucleotide
c 956	14	0.9	215	21	AAC31713	Human secreted pro
c 957	14	0.9	216	21	AAA59491	DNA encoding Smad3
c 958	14	0.9	218	21	AAC10055	Human secreted pro
c 959	14	0.9	220	15	AAQ76853	Human genome fragm
c 960	14	0.9	224	18	AAV78951	Staphylococcus aur

961	14	0.9	226	21	AAC81546	Human death domain
c 962	14	0.9	227	21	AAC93801	Cat flea hindgut a
c 963	14	0.9	227	21	AAC53782	Arabidopsis thalia
964	14	0.9	228	19	AAV61471	DR5-related clone
965	14	0.9	231	19	AAX11155	Human biallelic po
966	14	0.9	231	19	AAX11156	Human biallelic po
c 967	14	0.9	232	21	AAC05681	Human secreted pro
c 968	14	0.9	235	21	AAC28181	Human secreted pro
c 969	14	0.9	237	21	AAA63170	Ii reverse gene co
970	14	0.9	244	15	AAQ77374	Human genome fragm
971	14	0.9	247	21	AAC23267	Human secreted pro
c 972	14	0.9	249	18	AAV78833	Staphylococcus aur
973	14	0.9	251	19	AAX11670	Human biallelic po
c 974	14	0.9	251	21	AAC03407	Human secreted pro
975	14	0.9	253	21	AAC17903	Human secreted pro
c 976	14	0.9	254	15	AAQ71175	Merlin gene fragme
977	14	0.9	255	19	AAX11671	Human biallelic po
978	14	0.9	255	19	AAX12840	Human biallelic po
c 979	14	0.9	255	21	AAC00745	Human secreted pro
980	14	0.9	257	20	AAX54636	Human IL-5 antisen
981	14	0.9	257	20	AAV87644	EST clone EC172.
982	14	0.9	257	21	AAF20205	Human IL5 polynucl
c 983	14	0.9	257	21	AAC30906	Human secreted pro
984	14	0.9	257	21	AAA34083	Human adenosine re
c 985	14	0.9	258	18	AAV78906	Staphylococcus aur
c 986	14	0.9	260	21	AAC56220	Pinus radiata tran
c 987	14	0.9	260	21	AAC57111	Pinus radiata tran
988	14	0.9	261	21	AAZ45693	DNA encoding a fus
c 989	14	0.9	263	16	AAT23329	Human gene signatu
c 990	14	0.9	263	21	AAC12015	Human secreted pro
991	14	0.9	265	21	AAC71007	Single nucleotide
992	14	0.9	265	21	AAC09614	Human secreted pro
993	14	0.9	265	21	AAC27852	Human secreted pro
994	14	0.9	274	20	AAX51934	Human secreted pro
995	14	0.9	277	22	AAF68292	Human lung tumour
996	14	0.9	278	17	AAT09219	Virulence factor s
997	14	0.9	278	20	AAV86674	EST clone BG267.
c 998	14	0.9	278	21	AAA00066	Human colon cancer
999	14	0.9	279	20	AAV87257	EST clone BN422.
c1000	14	0.9	281	21	AAC10541	Human secreted pro

ALIGNMENTS

```
RESULT   1
AAA75359/c
ID   AAA75359 standard; DNA; 400 BP.
XX
AC   AAA75359;
XX
DT   15-JAN-2001  (first entry)
XX
DE   Fragment of IRES region of a COLIA1 targeted vector.
XX
KW   Somatic cell; nuclear transfer; genetic targeting; transgenic embryo;
KW   autologous transplantation; promoter trap gene targeting vector; ss.
XX
OS   Synthetic.
XX
PN   WO200051424-A2.
XX
PD   08-SEP-2000.
XX
PF   03-MAR-2000; 2000WO-GB00778.
XX
PR   04-MAR-1999;   99GB-0005033.
PR   20-JUL-1999;   99GB-0017023.
PR   30-DEC-1999;   99US-0475674.
XX
PA   (PPLT-) PPL THERAPEUTICS SCOTLAND LTD.
XX
PI   Colman A,  Schnieke AE,  Kind AJ,  Ayares DL,  Dai Y;
XX
DR   WPI; 2000-587219/55.
XX
PT   Genetic modification of somatic cell useful for nuclear transfer and
PT   transgenic animal production, involves genetic targeting event -
XX
PS   Example 1; Fig 2; 137pp; English.
XX
CC   The specification describes a method for preparing a somatic cell
CC   for nuclear transfer. The method involves modifying the genetic material
CC   of the somatic cell, while optionally maintaining the totipotency of
CC   the cell, by a genetic targeting event. The method is useful for nuclear
CC   transfer of genetic material from a somatic cell to a recipient cell to
CC   produce an animal embryo and totipotent or pluripotent cloned cell
CC   population. Nuclear transfer is also useful for producing a transgenic
CC   embryo or a transgenic foetus. The animal produced from such cells
CC   is useful for testing genetic changes due to the location of gene
CC   targeting. The method is also useful to validate loci in animal
CC   chromosomes which are suitable sites for transgene addition to cells.
CC   The method is also useful for cells that are required to have no
CC   unanticipated detrimental effects such as cells destained for autologous
CC   transplantation. The present sequence represents a portion of DNA
CC   sequence spanning the junction between the IRES neo gene and the
CC   5' region of shared homology. The sequence was derived from gene
CC   targeting of the neo gene to the COLIA1 locus in primary ovine foetal
CC   fibroblats cells, using the method of the invention.
XX
SQ   Sequence 400 BP; 90 A; 133 C; 83 G; 94 T; 0 other;

  Query Match             1.4%;  Score 21;  DB 21;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 0.99;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     603 ggattttccaaagtccatgtg 623
           |||||||||||||||||||||
Db     171 GGATTTTCCAAAGTCCATGTG 151


RESULT   2
AAV52471/c
ID   AAV52471 standard; DNA; 683 BP.
XX
AC   AAV52471;
XX
DT   23-OCT-1998  (first entry)
XX
DE   Streptococcus pneumoniae genome fragment SEQ ID NO:338.
XX
KW   Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KW   computer readable medium; vaccine; pharmaceutical composition; ds.
XX
OS   Streptococcus pneumoniae.
XX
PN   WO9818931-A2.
XX
PD   07-MAY-1998.
XX
PF   30-OCT-1997;   97WO-US19588.
XX
PR   31-OCT-1996;   96US-0029960.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Choi GH,  Dillon PJ,  Dougherty BA,  Fannon M;
PI   Kunsch CA,  Rosen CA;
XX
DR   WPI; 1998-272225/24.
XX
PT   Computer-readable medium with recorded Streptococcus pneumoniae
PT   polynucleotide sequences - useful in diagnostic kits and assays, and
PT   pharmaceutical compositions and vaccines for Streptococcus
PT   pneumoniae
```

XX
PS Claim 1; Page 1359-1360; 1409pp; English.
XX
CC The present invention describes a computer readable medium which has
CC the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)
CC recorded on it, or a representative fragment or a sequence at least 95%
CC identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in
CC SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
CC Streptococcus pneumoniae. The present invention also describes an
CC isolated nucleic acid molecule encoding a homologue of any of the
CC fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the
CC nucleic acid molecule is produced by a process comprising: (a) screening
CC a genomic DNA library using as a probe a target sequence defined by any
CC of the sequences in SEQ ID NO:1 to 391, identifying members of the
CC library which contain sequences that hybridise to the target sequence and
CC isolating the nucleic acid molecules from the members; or (b) isolating
CC mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
CC molecules whose nucleotide sequence is homologous to amplification
CC primers derived from the fragment of the S. pneumoniae genome to prime
CC the amplification and isolating the amplified sequences. The computer
CC readable medium can be used in a computer-based system for identifying
CC fragments of the S. pneumoniae genome of commercial importance, or
CC expression modulating fragments of the S. pneumoniae genome. Products
CC from the present invention can be used in diagnosis kits and assays, and
CC pharmaceutical compositions and vaccines for S. pneumoniae.
XX
SQ Sequence 683 BP; 181 A; 136 C; 141 G; 225 T; 0 other;

Query Match 1.3%; Score 20; DB 19; Length 683;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 cttttaatttttacgaacaa 56
||||||||||||||||||||
Db 46 CTTTTAATTTTTACGAACAA 27

RESULT 3
AAA69926
ID AAA69926 standard; cDNA; 262 BP.
XX
AC AAA69926;
XX
DT 07-NOV-2000 (first entry)
XX
DE Human ovarian carcinoma antigen polynucleotide SEQ ID NO:236.
XX
KW Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;
KW tumour antigen; identification; cytostatic; gene therapy; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200036107-A2.
XX
PD 22-JUN-2000.
XX
PF 17-DEC-1999; 99WO-US30270.
XX
PR 17-DEC-1998; 98US-0215681.
PR 17-DEC-1998; 98US-0216003.
PR 23-JUN-1999; 99US-0338933.
PR 24-SEP-1999; 99US-0404879.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Frudakis TN;
XX
DR WPI; 2000-431589/37.
XX
PT Immunogenic portion of an ovarian carcinoma protein and the nucleic
PT acid encoding it, useful for the diagnosis, prevention and treatment of
PT cancer, preferably ovarian cancer -
XX
PS Claim 18; Fig 15; 299pp; English.
XX
CC The present invention describes an isolated polypeptide comprising an
CC immunogenic portion of an ovarian carcinoma protein (or its variants).
CC Ovarian carcinoma proteins, and polynucleotides encoding them, have
CC cytostatic activity and can be used in gene therapy and vaccines.
CC Ovarian carcinoma polypeptides, nucleic acids, antibodies and vaccines
CC are useful for the prevention, diagnosis and treatment of cancer,
CC preferably ovarian cancer. AAA69691 to AAA70077 and AAB12552 to AAB12557
CC represent human ovarian carcinoma polynucleotides and proteins used in
CC the exemplification of the present invention.
XX
SQ Sequence 262 BP; 59 A; 38 C; 89 G; 76 T; 0 other;

Query Match 1.2%; Score 19; DB 21; Length 262;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 605 attttccaaagtccatgtg 623
|||||||||||||||||||
Db 100 attttccaaagtccatgtg 118

RESULT 4
AAA78188/c
ID AAA78188 standard; cDNA; 282 BP.
XX
AC AAA78188;
XX
DT 14-NOV-2000 (first entry)
XX
DE cDNA encoding human colon tumour polypeptide, SEQ ID NO:475.
XX
KW Human colon tumour polypeptide; tumour antigen; cancer; vaccine;
KW immunotherapy; diagnosis; progression; ss.
XX
OS Homo sapiens.
XX
PN WO200037643-A2.
XX
PD 29-JUN-2000.
XX
PF 23-DEC-1999; 99WO-US30909.
XX
PR 23-DEC-1998; 98US-0221298.
PR 02-JUL-1999; 99US-0347496.
PR 22-SEP-1999; 99US-0401064.
PR 19-NOV-1999; 99US-0444242.
PR 02-DEC-1999; 99US-0454150.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk J;
PI Wang T, Yuqiu J;
XX
DR WPI; 2000-442671/38.
XX
PT New colon tumor polypeptides used to inhibit the development of cancer,
PT especially colon cancer, and for diagnosing and monitoring the
PT progression of the cancer -
XX
PS Claim 29; Page 226; 229pp; English.
XX
CC Sequences AAA77722-A78199 represent 478 cDNAs encoding proteins or
CC portions of proteins which are associated with human colon tumours.
CC The invention also specifically discloses 8 human colon tumour proteins
CC (AAB11897-B11904). The nucleic acids, the polypeptides they encode, and
CC antigen presenting cells (APCs, preferably dendritic cells) expressing
CC such polypeptides may be used in vaccines that target tumour cells,
CC especially colon tumour cells, thereby inhibiting the development of
CC cancer. T-cells specific for the polypeptide expressed by the APC are

CC used to remove tumour cells from biological samples, especially blood or
CC fractions thereof. The sample or the isolated T-cells specific for the
CC polypeptide can then be used to inhibit cancer development. CD4+ and/or
CC CD8+ T-cells from a patient may be incubated with a polypeptide or
CC nucleic acid of the invention, or an APC expressing such a polypeptide,
CC to cause the proliferation of specific T-cells. The T-cells can be
CC cloned and then administered back to the patient to inhibit cancer
CC development. Nucleic acids encoding the polypeptides and antibodies
CC against the polypeptides may be used to determine the expression level
CC of a tumour protein of the invention, and therefore to determine whether
CC cancer cells are present. Such diagnostic methods may also be used to
CC monitor the progression of a cancer by repeating the processes at time
CC intervals, and comparing the current result to previous results. The
CC present sequence represents a cDNA encoding a human colon tumour
CC polypeptide.
XX
SQ Sequence 282 BP; 92 A; 89 C; 38 G; 63 T; 0 other;

Query Match 1.2%; Score 19; DB 21; Length 282;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 605 attttccaaagtccatgtg 623
||||||||||||||||||||
Db 167 ATTTTCCAAAGTCCATGTG 149

RESULT 5
AAA69935/c
ID AAA69935 standard; cDNA; 307 BP.
XX
AC AAA69935;
XX
DT 07-NOV-2000 (first entry)
XX
DE Human ovarian carcinoma antigen polynucleotide SEQ ID NO:245.
XX
KW Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;
KW tumour antigen; identification; cytostatic; gene therapy; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200036107-A2.
XX
PD 22-JUN-2000.
XX
PF 17-DEC-1999; 99WO-US30270.
XX
PR 17-DEC-1998; 98US-0215681.
PR 17-DEC-1998; 98US-0216003.
PR 23-JUN-1999; 99US-0338933.
PR 24-SEP-1999; 99US-0404879.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Frudakis TN;
XX
DR WPI; 2000-431589/37.
XX
PT Immunogenic portion of an ovarian carcinoma protein and the nucleic
PT acid encoding it, useful for the diagnosis, prevention and treatment of
PT cancer, preferably ovarian cancer -
XX
PS Claim 18; Fig 15; 299pp; English.
XX
CC The present invention describes an isolated polypeptide comprising an
CC immunogenic portion of an ovarian carcinoma protein (or its variants).
CC Ovarian carcinoma proteins, and polynucleotides encoding them, have
CC cytostatic activity and can be used in gene therapy and vaccines.
CC Ovarian carcinoma polypeptides, nucleic acids, antibodies and vaccines
CC are useful for the prevention, diagnosis and treatment of cancer,
CC preferably ovarian cancer. AAA69691 to AAA70077 and AAB12552 to AAB12557
CC represent human ovarian carcinoma polynucleotides and proteins used in
CC the exemplification of the present invention.
XX
SQ Sequence 307 BP; 83 A; 102 C; 55 G; 67 T; 0 other;

Query Match 1.2%; Score 19; DB 21; Length 307;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 605 attttccaaagtccatgtg 623
||||||||||||||||||||
Db 208 ATTTTCCAAAGTCCATGTG 190

RESULT 6
AAT80189/c
ID AAT80189 standard; cDNA; 977 BP.
XX
AC AAT80189;
XX
DT 01-APR-1998 (first entry)
XX
DE Human collagen 1A1 coding sequence containing 3210A polymorphism.
XX
KW Collagen; human; expression suppression; autosomal dominant disease;
KW suppression effector; polygenic disease; therapy; manic depression;
KW osteogenesis imperfecta; rhodopsin-linked retinitis pigmentosa; glaucoma;
KW age-related macular degeneration; disease-causing mutation;
KW inherited cancer; polymorphism targeting; ss.
XX
OS Homo sapiens.
XX
PN WO9732024-A1.
XX
PD 04-SEP-1997.
XX
PF 03-MAR-1997; 97WO-GB00574.
XX
PR 01-MAR-1996; 96GB-0004449.
XX
PA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX
PI Farrar GJ, Humphries P, Kenna PF;
XX
DR WPI; 1997-448692/41.
XX
PT Suppression of one allele of an endogenous gene - used for treating
PT autosomal dominant or polygenic diseases, or to increase effect of
PT an administered agent
XX
PS Example 2; Fig 17; 83pp; English.
XX
CC This sequence represents the collagen 1A1 coding sequence containing the
CC A polymorphism at 3210. Collagen is an example of a gene that can be used
CC in the method of the invention. The method of the invention is a novel
CC strategy for suppressing expression of one allele of an endogenous gene,
CC comprises treating the gene with suppression effectors that bind to a
CC polymorphism within, or close to, the gene so that the allele is
CC suppressed, exclusively or preferentially. The method is used in the
CC treatment of autosomal dominant or polygenic diseases, also to increase
CC activity of an administered compound, particularly where the effect of
CC this compound is inhibited by the allele or its gene product. Typical
CC applications are in the treatment of rhodopsin-linked retinitis
CC pigmentosa and osteogenesis imperfecta, or in age-related macular
CC degeneration, glaucoma, manic depression, cancers with an inherited
CC component etc. The same suppression effectors can be used to suppress
CC many different disease-causing mutations in a single gene, eliminating
CC the need for a specific therapy for each mutation. Targeting of
CC polymorphisms, rather than the disease-related allele itself, allows a
CC more flexible choice of target sequence.
XX
SQ Sequence 977 BP; 221 A; 253 C; 202 G; 227 T; 74 other;

```
  Query Match             1.2%;  Score 19;  DB 18;  Length 977;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     605 attttccaaagtccatgtg 623
           ||||||||||||||||||
Db     388 ATTTTCCAAAGTCCATGTG 370


RESULT   7
AAC77648/c
ID   AAC77648 standard; cDNA; 2192 BP.
XX
AC   AAC77648;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated gene sequence SEQ ID NO:42.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   P-PSDB; AAB43439.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 1; Page 640-641; 2352pp; English.
XX
CC   AAC77607 to AAC78448 encode the human cancer associated proteins given
CC   in AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells, to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reactions, graft versus host disease and organ
CC   rejection, modulate haemostatic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
CC   bacterial or viral infections. The peptides, nucleotides, antibodies,
CC   agonists and antagonists may be also be used in drug screens. AAC78449 to
CC   AAC78457 and AAB44240 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 2192 BP; 419 A; 723 C; 620 G; 429 T; 1 other;


  Query Match             1.2%;  Score 19;  DB 21;  Length 2192;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     605 attttccaaagtccatgtg 623
           ||||||||||||||||||
Db    2049 ATTTTCCAAAGTCCATGTG 2031


RESULT   8
AAN71302
ID   AAN71302 standard; DNA; 3871 BP.
XX
AC   AAN71302;
XX
DT   30-APR-1991  (first entry)
XX
DE   HSV-1 gB and surrounding regions.
XX
KW   Vaccine; prophylaxis; treatment; Herpes Simplex Virus-1;
KW   glycoprotein; gB; ss.
XX
OS   Herpes simplex virus type 1 (KOS).
XX
FH   Key             Location/Qualifiers
FT   misc_RNA        1..375
FT                   /*tag= a
FT                   /note= "5' extra sequences beginning with the XhoI
FT                        site"
FT   CAAT_signal     406..410
FT                   /*tag= b
FT                   /number= 1
FT   CAAT_signal     443..448
FT                   /*tag= c
FT                   /number= 2
FT   TATA_signal     476..479
FT                   /*tag= d
FT   misc_RNA        501..789
FT                   /*tag= e
FT                   /label= mRNA start sequence
FT                   /note= "501 is possible start site"
FT   misc_RNA        504..789
FT                   /*tag= f
FT                   /label= mRNA start sequence
FT                   /note= "504 is possible start site"
FT   misc_RNA        506..789
FT                   /*tag= g
FT                   /label= mRNA start sequence
FT                   /note= "506 is possible start site"
FT   CDS             790..3498
FT                   /*tag= h
FT                   /label= HSV-1 gB
FT                   /note= "includes N-terminal hydrophobic leader and
FT                        a membrane-spanning sequence, a C-terminal
FT                        ionic sequence, and 9 N-linked
FT                        saccharide-addition sites"
FT   3'UTR           3499..3549
FT                   /*tag= i
FT   polyA_signal    3518..3525
FT                   /*tag= j
FT   polyA_site      3549..3549
FT                   /*tag= k
FT   misc_RNA        3549..3997
FT                   /*tag= l
FT                   /note= "3' nonessential sequences to the BamHI
FT                        site"
XX
```

```
PN   US4642333-A.
XX
PD   10-FEB-1987.
XX
PF   20-JUN-1984;   84US-0622496.
XX
PR   20-JUN-1985;   85US-0622496.
PR   16-SEP-1983;   83US-0532996.
XX
PA   (PERS/) PERSON S.
XX
PI   Person S;
XX
DR   WPI; 1987-056354/08.
XX
PT   Amino acid chain of glyco:protein B of HSV-1 and 2 - prepd. as
PT   recombinant and used for vaccines for herpes simplex virus types 1
PT   and 2.
XX
PS   Example; Table 1; 16pp; English.
XX
CC   !NOTE!  This sequence has been indexed as represented in the
CC   specification, except that bases 'E' have been replaced by 'N'.
CC   The features have been indexed according to the legend of table 1 on
CC   column 19/20 and the Sequence Summary of column 5 (sic).  Note that
CC   the base numbering of the features does not correspond to the
CC   the sequence numbering below.
CC   For another DNA sequence of HSB-1 gB see AAN71303 (AAP71135),
CC   and for HSV-2 gB see AAN71399 (AAP71136).
CC   A pure non-glycosylated amino acid (AA) chain comprising a sequence
CC   corresponding to that occurring in glycoprotein B of HSV-1 or HSV-2
CC   virus which is antigenic to HSV-1 of HSV-2, which contains no more
CC   than 750 AA residues, and which includes AA residues 135-649
CC   inclusive is claimed. It can be used to produce vaccines for
CC   prophylaxis and treatment of HSV-1 and HSV-2.
XX
SQ   Sequence 3871 BP; 743 A; 1402 C; 544 G; 754 T; 428 other;


  Query Match             1.2%;  Score 19;  DB 8;  Length 3871;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     325 gaggagctcgcccggttca 343
           ||||||||||||||||||||
Db     364 gaggagctcgcccggttca 382


RESULT   9
AAN50516
ID   AAN50516 standard; DNA; 3996 BP.
XX
AC   AAN50516;
XX
DT   21-NOV-1991  (first entry)
XX
DE   Sequence encoding the glycoprotein B (gB) of Herpes simplex virus
DE   type 1 (HSV-1).
XX
KW   Vaccine; glycoprotein B; ss.
XX
OS   Herpes simplex virus type 1.
XX
FH   Key             Location/Qualifiers
FT   CDS             790..3501
FT                   /*tag=  a
FT   mRNA            501
FT                   /*tag=  b
FT   mRNA            504
FT                   /*tag=  c
FT   mRNA            506
FT                   /*tag=  d
FT   misc_signal     476..482
FT                   /*tag=  e
FT   misc_signal     3518..3523
FT                   /*tag=  f
FT   polyA_site      3543
FT                   /*tag=  g
XX
PN   EP133063-A.
XX
PD   13-FEB-1985.
XX
PF   22-JUN-1984;   84EP-0401312.
XX
PR   16-SEP-1983;   83US-0532996.
PR   23-JUN-1983;   83US-0506986.
PR   20-JUN-1984;   84US-0622496.
XX
PA   (PERS/) PERSON S.
XX
PI   Person S;
XX
DR   WPI; 1985-039636/07.
XX
PT   New amino acid chains of glyco-protein B of herpes simplex virus
PT   - prepd. by recombination DNA methods for use in vaccines
XX
PS   Disclosure; Table 1, Page 34-35; 40pp; English.
XX
CC   The inventors claim a non-glycosylated amino acid chain comprising a
CC   sequence corresponding to that occurring in glycoprotein B of HSV-1
CC   or HSV-2. Preferably, when the chain comprises a sequence corresp.
CC   to that occurring in glycoprotein B of HSV-2, it contains 4-750 AA
CC   residues. The chain esp.has a molecular wt. of 65000 daltons and
CC   includes AA residues 165-629 from the HSV-2 glycoprotein B sequence.
XX
SQ   Sequence 3996 BP; 745 A; 1376 C; 1207 G; 668 T; 0 other;


  Query Match             1.2%;  Score 19;  DB 6;  Length 3996;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     325 gaggagctcgcccggttca 343
           ||||||||||||||||||||
Db     348 gaggagctcgcccggttca 366


RESULT  10
AAT20801/c
ID   AAT20801 standard; cDNA to mRNA; 430 BP.
XX
AC   AAT20801;
XX
DT   09-JUL-1996  (first entry)
XX
DE   Human gene signature HUMGS02049.
XX
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
XX
OS   Homo sapiens.
XX
PN   WO9514772-A1.
XX
PD   01-JUN-1995.
XX
PF   11-NOV-1994;   94WO-JP01916.
XX
PR   12-NOV-1993;   93JP-0355504.
XX
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
XX
```

```
PI   Matsubara K,  Okubo K;
XX
DR   WPI; 1995-206931/27.
XX
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
XX
PS   Claim 1; Page 727; 2245pp; Japanese.
XX
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in AAT19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
XX
SQ   Sequence 430 BP; 112 A; 128 C; 82 G; 86 T; 22 other;


  Query Match             1.2%;  Score 18;  DB 16;  Length 430;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    606 ttttccaaagtccatgtg 623
          ||||||||||||||||||
Db    337 TTTTCCAAAGTCCATGTG 320


RESULT  11
AAZ29568/c
ID   AAZ29568 standard; cDNA; 528 BP.
XX
AC   AAZ29568;
XX
DT   22-MAR-2000  (first entry)
XX
DE   Rat cDNA clone rC1R11.
XX
KW   Clone rC1R11; retinoic acid; A2M; Alpha-2-macroglobulin;
KW   immunoaffinity probe; cell proliferation; cell differentiation;
KW   alveolar formation; pulmonary disorder; emphysema;
KW   immature lung development; ds.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   polyA_signal    495..500
FT                   /*tag=  a
XX
PN   WO9964057-A1.
XX
PD   16-DEC-1999.
XX
PF   11-JUN-1999;   99WO-US13287.
XX
PR   11-JUN-1998;   98US-0088977.
PR   11-JUN-1999;   99US-0088977.
XX
PA   (GEOU ) UNIV GEORGETOWN MEDICAL CENT.
XX
PI   Chan W;
XX
DR   WPI; 2000-116454/10.
XX
PT   Novel nucleic acid sequences, polypeptides encoded by them in treating
PT   emphysema in infants and for screening compounds promoting alveolar
PT   formation  -
XX
PS   Claim 1; Page 22; 67pp; English.
XX
CC   The present sequence is a cDNA clone rC1R11, isolated from rat lung
CC   tissue. Its expression is induced by retinoic acid in elastase treated
CC   rat lung and is co-localized with emphysema due to A2M
CC   (alpha-2-macroglobulin) deficiency. Polypeptide encoded by the clone can
CC   be used as immunoaffinity probe to identify related cDNAs from cDNA
CC   expression library prepared from human or other species lung tissue
CC   and to screen compound libraries for molecules that agonise or
CC   antagonise retinoic acid induced cell proliferation or differentiation.
CC   It is involved in cell proliferation/differentiation and alveolar
CC   formation and can be used to treat pulmonary disorders like emphysema and
CC   immature lung development in premature infants.
XX
SQ   Sequence 528 BP; 149 A; 101 C; 104 G; 174 T; 0 other;


  Query Match             1.2%;  Score 18;  DB 21;  Length 528;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    459 taactatttcttagaaga 476
          ||||||||||||||||||
Db     91 TAACTATTTCTTAGAAGA 74


RESULT  12
AAZ34777/c
ID   AAZ34777 standard; cDNA; 546 BP.
XX
AC   AAZ34777;
XX
DT   15-FEB-2000  (first entry)
XX
DE   Rice sucrose phosphate synthase cDNA.
XX
KW   Sucrose phosphate synthase; SPS; transgenic plant; rice; ss.
XX
OS   Oryza sativa.
XX
FH   Key             Location/Qualifiers
FT   CDS             8..343
FT                   /*tag=  a
FT                   /partial
XX
PN   WO9957285-A2.
XX
PD   11-NOV-1999.
XX
PF   06-MAY-1999;   99WO-US09865.
XX
PR   07-MAY-1998;   98US-0084529.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Orozco EM,  Caimi PG,  Weng Z,  Tarczynski MC;
XX
DR   WPI; 2000-052973/04.
DR   P-PSDB; AAY32270.
XX
PT   Gene encoding sucrose biosynthetic enzyme useful for producing
PT   transgenic plant with altered level of enzyme activity  -
XX
PS   Claim 8; Page 43; 62pp; English.
XX
CC   This nucleotide sequence comprises a portion of the cDNA insert in
CC   clone r10n.pk086.123 encoding a substantial portion of rice sucrose
```

```
CC   phosphate synthase (SPS, EC-2.4.1.14). The clone was isolated from
CC   a rice 15-day-old leaf cDNA library. The invention provides isolated
CC   nucleic acid fragments (see AAZ34772-82) encoding plant SPS (see
CC   AAY32265-75). The invention also relates to the construction of a
CC   chimeric gene encoding all or a portion of the SPS, in sense or
CC   antisense orientation, such that expression of the chimeric gene
CC   results in production of altered levels of the SPS in a transformed
CC   host cell. Thus, SPS nucleic acids are useful for creating
CC   transgenic plants in which the SPS is present in higher or lower
CC   levels than normal, thereby altering the level of sucrose
CC   biosynthetic activity in the cells. They are also used for
CC   production of SPS in heterologous microbial host cells, and for the
CC   design of probes and primers.
XX
SQ   Sequence 546 BP; 167 A; 113 C; 115 G; 147 T; 4 other;


  Query Match             1.2%;  Score 18;  DB 21;  Length 546;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     681 tataacaatgatatgctt 698
           ||||||||||||||||||
Db     247 TATAACAATGATATGCTT 230


RESULT  13
AAC66433
ID   AAC66433 standard; DNA; 930 BP.
XX
AC   AAC66433;
XX
DT   14-FEB-2001  (first entry)
XX
DE   Human secreted protein coding sequence SEQ ID NO: 34.
XX
KW   Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW   antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW   vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW   cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; human; secreted protein; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058350-A1.
XX
PD   05-OCT-2000.
XX
PF   22-MAR-2000; 2000WO-US07483.
XX
PR   26-MAR-1999;   99US-0126596.
PR   22-DEC-1999;   99US-0171552.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM,  Komatsoulis G;
XX
DR   WPI; 2000-602357/57.
DR   P-PSDB; AAB32025.
XX
PT   Nucleic acid molecules encoding human secreted proteins, used in
PT   preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT   Parkinson's diseases and cancers -
XX
PS   Claim 1; Page 347; 423pp; English.
XX
CC   The invention relates to the isolation of genes AAC66410-C66458 encoding
CC   49 human secreted proteins AAB32002-B32050. The genes can be used to
CC   generate fusion proteins by linking to the gene for the human
CC   immunoglobulin G Fc portion (SEQID1) for increasing the stability of
CC   the fusion protein as compared to the human protein only. The genes and
CC   proteins are useful for preventing, ameliorating or treating medical
CC   conditions, e.g. by protein or gene therapy. The genes are isolated
CC   from a range of human tissues disclosed in the specification. The
CC   nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC   and ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC   (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC   haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC   colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
CC   wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC   epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC   and parasitic infections.
XX
SQ   Sequence 930 BP; 284 A; 150 C; 173 G; 323 T; 0 other;


  Query Match             1.2%;  Score 18;  DB 21;  Length 930;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     554 tcttcatggaaaaacagg 571
           ||||||||||||||||||
Db      50 tcttcatggaaaaacagg 67


RESULT  14
AAC86928
ID   AAC86928 standard; DNA; 1149 BP.
XX
AC   AAC86928;
XX
DT   02-APR-2001  (first entry)
XX
DE   DNA encoding nuclear-encoded mitochondrial elongation factor Ts.
XX
KW   Mitochondrial elongation factor Ts; EF-Ts; leimyosarcoma; restenosis;
KW   smooth muscle cell proliferation; translational elongation factor;
KW   hyperplasia; cancer; angioplasty; unstable angina; atheromatous plaque;
KW   atherosclerosis; tumour; ss.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   CDS             8..982
FT                   /*tag=  a
FT                   /product= "nuclear-encoded mitochondrial elongation
FT                        factor Ts"
FT                   /transl_except= (979..981, aa: Xaa)
FT                   /note= "Xaa represents Glu and Glx"
FT   polyA_signal    1116..1121
FT                   /*tag=  b
XX
PN   FR2793798-A1.
XX
PD   24-NOV-2000.
XX
PF   21-MAY-1999;   99FR-0006530.
XX
PR   21-MAY-1999;   99FR-0006530.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   McGregor JL,  Bourdillon MC,  Zibara K;
XX
DR   WPI; 2001-064130/08.
DR   P-PSDB; AAB31160.
XX
PT   Antisense oligonucleotide, useful for treating e.g. cancer or
PT   restenosis, is directed against the translational elongation factor
PT   gene -
XX
PS   Claim 4; Fig 5A; 48pp; French.
XX
```

CC The present sequence encodes a rat nuclear-encoded mitochondrial
CC elongation factor Ts (EF-Ts). EF-Ts is overexpressed in proliferating
CC smooth muscle cells. The specification describes an oligonucleotide
CC of 8-100 nucleotides that hybridises with nucleic acid encoding the
CC translational elongation factor (EF-Ts). This oligonucleotide inhibits
CC expression of EF-Ts in mammals. The oligonucleotide is used to treat
CC hyperplasia, particularly cancer (especially leimyosarcoma),
CC atherosclerosis, restenosis caused by angioplasty and unstable angina.
CC The oligonucleotide can be used as a probe or primer to stage
CC atheromatous plaque, to study polymorphisms of the gene for (EF-Ts),
CC particularly in relation to hyperplasia, and to detect tumours.
XX
SQ Sequence 1149 BP; 271 A; 288 C; 322 G; 268 T; 0 other;

Query Match 1.2%; Score 18; DB 22; Length 1149;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 651 catatctcctggatcctt 668
||||||||||||||||||
Db 864 catatctcctggatcctt 881

RESULT 15
AAA70248
ID AAA70248 standard; DNA; 1689 BP.
XX
AC AAA70248;
XX
DT 07-NOV-2000 (first entry)
XX
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:381.
XX
KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KW antimalarial; malaria; protozoacide; infection; insecticide; ds.
XX
OS Plasmodium falciparum.
XX
PN WO200025728-A2.
XX
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US26796.
XX
PR 05-NOV-1998; 98US-0107131.
XX
PA (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX
DR WPI; 2000-365347/31.
XX
PT Proteins encoded by chromosome 2 of the human malarial parasite,
PT Plasmodium falciparum, useful as antimalarial vaccines and in the
PT diagnosis of P.falciparum infection -
XX
PS Disclosure; Page 559-560; 577pp; English.
XX
CC The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against P. falciparum infection comprising (I) or (II).
CC (I) and (II) are useful for the development of vaccines against
CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal
CC antibody raised to immunogens comprising the sequences of (I), are
CC useful in the detection of infection with P. falciparum. Furthermore,
CC (I) (especially when they are rifins or secreted or membrane proteins)
CC can aid the identification of drugs to treat or prevent P. falciparum
CC infection, or they can be used to identify drug resistance in
CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
CC subsequent identification of proteins encoded by it will help to expand
CC our understanding of parasite biology, a process hampered by the
CC complexity of the parasitic lifecycle, and provide new targets for
CC vaccine and drug development. Parasite resistance to drugs and mosquito
CC resistance to insecticides have led to a resurgence of malaria in many
CC parts of the world, and there is a pressing need for vaccines and new
CC drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide
CC and protein sequences given in the present invention, but which are not
CC specifically mentioned within the specification.
XX
SQ Sequence 1689 BP; 768 A; 195 C; 228 G; 498 T; 0 other;

Query Match 1.2%; Score 18; DB 21; Length 1689;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 805 atactagatcatttaaat 822
||||||||||||||||||
Db 853 atactagatcatttaaat 870

Search completed: October 30, 2001, 13:05:42
Job time: 3469 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2001, 11:43:48 ; Search time 55.04 Seconds
(without alignments)
5283.101 Million cell updates/sec

Title: US-09-619-049-854
Perfect score: 1536
Sequence: 1 atgaggcttccttgggtgat.....ggactctaatgatggattct 1536

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 324599 seqs, 94655562 residues

Word size : 0

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	17	1.1	303	2	US-08-637-759B-24	Sequence 24, Appl
	2	17	1.1	303	3	US-08-871-355A-24	Sequence 24, Appl
c	3	17	1.1	5746	2	US-08-637-759B-38	Sequence 38, Appl
c	4	17	1.1	5746	3	US-08-871-355A-38	Sequence 38, Appl
c	5	17	1.1	9278	1	US-08-243-542-9	Sequence 9, Appli
c	6	17	1.1	9278	1	US-08-477-407-9	Sequence 9, Appli
c	7	17	1.1	9278	1	US-08-484-355-9	Sequence 9, Appli
c	8	16	1.0	450	2	US-08-345-321-5	Sequence 5, Appli
	9	16	1.0	1178	2	US-08-491-988-25	Sequence 25, Appl
	10	16	1.0	1178	2	US-08-491-988-28	Sequence 28, Appl
	11	16	1.0	1202	2	US-08-491-988-27	Sequence 27, Appl
	12	16	1.0	1226	2	US-08-491-988-12	Sequence 12, Appl
	13	16	1.0	1235	2	US-08-491-988-11	Sequence 11, Appl
	14	16	1.0	1259	2	US-08-491-988-10	Sequence 10, Appl
	15	16	1.0	1259	2	US-08-491-988-24	Sequence 24, Appl
	16	16	1.0	1295	2	US-08-491-988-26	Sequence 26, Appl
c	17	16	1.0	1376	2	US-08-868-288A-2	Sequence 2, Appli
c	18	16	1.0	1376	3	US-09-235-373-2	Sequence 2, Appli
c	19	16	1.0	1376	3	US-09-388-993-2	Sequence 2, Appli
c	20	16	1.0	2210	1	US-07-710-361-2	Sequence 2, Appli
c	21	16	1.0	2223	1	US-07-684-135A-1	Sequence 1, Appli
c	22	16	1.0	2582	4	US-09-102-528-31	Sequence 31, Appl
	23	16	1.0	4138	1	US-08-323-474-1	Sequence 1, Appli
	24	16	1.0	4138	5	PCT-US93-06093-1	Sequence 1, Appli
c	25	16	1.0	5872	4	US-09-102-528-32	Sequence 32, Appl
	26	16	1.0	7587	4	US-08-378-313-22	Sequence 22, Appl
	27	16	1.0	7876	1	US-08-225-480-1	Sequence 1, Appli
	28	16	1.0	7876	5	PCT-US93-08115-1	Sequence 1, Appli
	29	15	1.0	24	1	US-08-179-738-26	Sequence 26, Appl
	30	15	1.0	24	2	US-08-628-145-26	Sequence 26, Appl
c	31	15	1.0	30	1	US-08-802-547-13	Sequence 13, Appl
	32	15	1.0	102	3	US-08-817-335-4	Sequence 4, Appli
c	33	15	1.0	135	2	US-08-611-757-7	Sequence 7, Appli
c	34	15	1.0	135	5	PCT-US95-05980-7	Sequence 7, Appli
	35	15	1.0	206	1	US-08-248-474-45	Sequence 45, Appl
	36	15	1.0	206	3	US-08-756-849-45	Sequence 45, Appl
	37	15	1.0	303	1	US-08-307-499-9	Sequence 9, Appli
	38	15	1.0	303	4	US-09-299-268-9	Sequence 9, Appli
	39	15	1.0	350	1	US-08-038-071-4	Sequence 4, Appli
c	40	15	1.0	434	1	US-07-872-678A-5	Sequence 5, Appli
	41	15	1.0	579	2	US-08-820-170A-5	Sequence 5, Appli
	42	15	1.0	579	3	US-09-055-699-5	Sequence 5, Appli
	43	15	1.0	579	4	US-09-273-565-5	Sequence 5, Appli
	44	15	1.0	767	1	US-08-184-604-1	Sequence 1, Appli
c	45	15	1.0	818	1	US-08-271-354-15	Sequence 15, Appl
c	46	15	1.0	818	2	US-08-565-861-15	Sequence 15, Appl
c	47	15	1.0	818	5	PCT-US94-07658-15	Sequence 15, Appl
c	48	15	1.0	825	1	US-08-307-499-48	Sequence 48, Appl
c	49	15	1.0	825	4	US-09-299-268-48	Sequence 48, Appl
	50	15	1.0	846	2	US-08-619-542B-43	Sequence 43, Appl
	51	15	1.0	1015	2	US-08-820-170A-6	Sequence 6, Appli
	52	15	1.0	1015	3	US-09-055-699-6	Sequence 6, Appli
	53	15	1.0	1015	4	US-09-273-565-6	Sequence 6, Appli
	54	15	1.0	1033	1	US-08-592-126-92	Sequence 92, Appl
c	55	15	1.0	1104	1	US-08-307-499-57	Sequence 57, Appl
c	56	15	1.0	1104	4	US-09-299-268-57	Sequence 57, Appl
c	57	15	1.0	1122	1	US-08-625-876-1	Sequence 1, Appli
c	58	15	1.0	1122	1	US-08-625-876-3	Sequence 3, Appli
c	59	15	1.0	1122	1	US-08-625-876-5	Sequence 5, Appli
c	60	15	1.0	1199	3	US-08-755-587-38	Sequence 38, Appl
	61	15	1.0	1207	2	US-08-525-864A-5	Sequence 5, Appli
c	62	15	1.0	1235	6	5171685-1	Patent No. 5171685
c	63	15	1.0	1236	6	5518916-1	Patent No. 5518916
c	64	15	1.0	1320	3	US-08-989-251-28	Sequence 28, Appl
c	65	15	1.0	1320	3	US-09-340-250-28	Sequence 28, Appl
c	66	15	1.0	1358	1	US-08-419-009-29	Sequence 29, Appl
	67	15	1.0	1380	1	US-08-599-252-82	Sequence 82, Appl
	68	15	1.0	1380	1	US-08-436-074-55	Sequence 55, Appl
	69	15	1.0	1380	5	PCT-US96-06352-82	Sequence 82, Appl
	70	15	1.0	1380	5	PCT-US96-06583-82	Sequence 82, Appl
c	71	15	1.0	1396	2	US-09-151-611-2	Sequence 2, Appli
	72	15	1.0	1399	4	US-08-617-785-5	Sequence 5, Appli
c	73	15	1.0	1426	1	US-08-419-009-28	Sequence 28, Appl
c	74	15	1.0	1435	1	US-08-419-009-26	Sequence 26, Appl
	75	15	1.0	1474	1	US-08-300-903A-10	Sequence 10, Appl
	76	15	1.0	1510	1	US-08-300-903A-12	Sequence 12, Appl
	77	15	1.0	1517	1	US-08-032-869A-1	Sequence 1, Appli
	78	15	1.0	1517	1	US-08-032-869A-2	Sequence 2, Appli
	79	15	1.0	1517	1	US-08-472-673-1	Sequence 1, Appli
	80	15	1.0	1517	1	US-08-472-673-2	Sequence 2, Appli
	81	15	1.0	1517	1	US-08-475-782-1	Sequence 1, Appli
	82	15	1.0	1517	1	US-08-475-782-2	Sequence 2, Appli
	83	15	1.0	1517	2	US-08-472-678-1	Sequence 1, Appli
	84	15	1.0	1517	2	US-08-472-678-2	Sequence 2, Appli
	85	15	1.0	1517	2	US-08-484-503-1	Sequence 1, Appli
	86	15	1.0	1517	2	US-08-484-503-2	Sequence 2, Appli
	87	15	1.0	1518	4	US-08-944-368A-1	Sequence 1, Appli
	88	15	1.0	1521	1	US-08-496-855A-3	Sequence 3, Appli
	89	15	1.0	1521	2	US-07-938-154-9	Sequence 9, Appli
	90	15	1.0	1521	5	PCT-US91-02311-9	Sequence 9, Appli
	91	15	1.0	1534	1	US-08-300-903A-6	Sequence 6, Appli
c	92	15	1.0	1583	3	US-08-976-255-8	Sequence 8, Appli
	93	15	1.0	1641	1	US-08-300-903A-8	Sequence 8, Appli
c	94	15	1.0	1650	1	US-08-354-240A-1	Sequence 1, Appli
c	95	15	1.0	1650	1	US-08-354-240A-3	Sequence 3, Appli
c	96	15	1.0	1653	1	US-08-354-240A-5	Sequence 5, Appli
c	97	15	1.0	1656	3	US-09-111-752-6	Sequence 6, Appli
c	98	15	1.0	1656	3	US-09-111-752-9	Sequence 9, Appli
	99	15	1.0	1680	2	US-08-700-152A-3	Sequence 3, Appli
c	100	15	1.0	1722	3	US-08-718-425-1	Sequence 1, Appli

c	101	15	1.0	1722	4	US-08-875-277A-1	Sequence 1, Appli
c	102	15	1.0	1747	1	US-08-176-620A-1	Sequence 1, Appli
c	103	15	1.0	1747	1	US-08-463-862-1	Sequence 1, Appli
c	104	15	1.0	1747	2	US-08-461-985-1	Sequence 1, Appli
c	105	15	1.0	1747	2	US-08-458-887-1	Sequence 1, Appli
c	106	15	1.0	1811	3	US-08-867-352-22	Sequence 22, Appl
c	107	15	1.0	1816	1	US-08-061-889-1	Sequence 1, Appli
c	108	15	1.0	1816	1	US-08-462-611-1	Sequence 1, Appli
c	109	15	1.0	1816	5	PCT-US94-05378-1	Sequence 1, Appli
c	110	15	1.0	1844	1	US-07-872-644-26	Sequence 26, Appl
c	111	15	1.0	1844	1	US-08-297-494-26	Sequence 26, Appl
c	112	15	1.0	1844	1	US-08-297-510-26	Sequence 26, Appl
c	113	15	1.0	1844	1	US-08-479-532-26	Sequence 26, Appl
c	114	15	1.0	1844	1	US-08-455-526-26	Sequence 26, Appl
c	115	15	1.0	1844	1	US-08-455-525-26	Sequence 26, Appl
c	116	15	1.0	1844	3	US-09-139-491-26	Sequence 26, Appl
c	117	15	1.0	1844	5	PCT-US92-03222-26	Sequence 26, Appl
c	118	15	1.0	1862	1	US-08-179-738-4	Sequence 4, Appli
c	119	15	1.0	1862	2	US-08-628-145-4	Sequence 4, Appli
	120	15	1.0	1907	3	US-08-771-986A-1	Sequence 1, Appli
	121	15	1.0	1907	3	US-08-769-802A-1	Sequence 1, Appli
c	122	15	1.0	1939	6	5198542-3	Patent No. 5198542
	123	15	1.0	2032	5	PCT-US95-07721-5	Sequence 5, Appli
c	124	15	1.0	2174	2	US-08-665-040-1	Sequence 1, Appli
	125	15	1.0	2205	4	US-08-687-590-59	Sequence 59, Appl
	126	15	1.0	2265	1	US-08-242-689-1	Sequence 1, Appli
	127	15	1.0	2290	1	US-08-584-226-1	Sequence 1, Appli
	128	15	1.0	2319	1	US-08-286-767-1	Sequence 1, Appli
	129	15	1.0	2320	1	US-08-090-526-1	Sequence 1, Appli
	130	15	1.0	2320	1	US-08-486-715-1	Sequence 1, Appli
	131	15	1.0	2320	1	US-08-486-719-1	Sequence 1, Appli
	132	15	1.0	2320	1	US-08-271-354-1	Sequence 1, Appli
	133	15	1.0	2320	1	US-08-476-100-1	Sequence 1, Appli
	134	15	1.0	2320	2	US-08-565-861-1	Sequence 1, Appli
	135	15	1.0	2320	3	US-08-475-749-1	Sequence 1, Appli
	136	15	1.0	2320	5	PCT-US94-07658-1	Sequence 1, Appli
c	137	15	1.0	2333	1	US-08-392-678-33	Sequence 33, Appl
c	138	15	1.0	2334	1	US-08-457-304A-33	Sequence 33, Appl
c	139	15	1.0	2334	1	US-08-456-701A-33	Sequence 33, Appl
c	140	15	1.0	2363	2	US-08-671-978A-11	Sequence 11, Appl
	141	15	1.0	2391	2	US-08-817-900-1	Sequence 1, Appli
	142	15	1.0	2400	4	US-07-861-458C-94	Sequence 94, Appl
	143	15	1.0	2431	3	US-08-985-526-35	Sequence 35, Appl
c	144	15	1.0	2445	1	US-08-122-520C-8	Sequence 8, Appli
	145	15	1.0	2450	2	US-08-466-589-9	Sequence 9, Appli
	146	15	1.0	2450	2	US-08-700-636-9	Sequence 9, Appli
	147	15	1.0	2450	3	US-08-467-574-9	Sequence 9, Appli
c	148	15	1.0	2465	3	US-08-619-812-5	Sequence 5, Appli
c	149	15	1.0	2781	3	US-08-749-522-4	Sequence 4, Appli
c	150	15	1.0	2964	2	US-08-846-790A-2	Sequence 2, Appli
c	151	15	1.0	2964	3	US-08-935-333-2	Sequence 2, Appli
	152	15	1.0	2998	2	US-08-841-349-12	Sequence 12, Appl
	153	15	1.0	3344	1	US-07-718-575-13	Sequence 13, Appl
	154	15	1.0	3344	1	US-08-481-206-13	Sequence 13, Appl
	155	15	1.0	3344	2	US-08-486-269A-13	Sequence 13, Appl
	156	15	1.0	3441	2	US-08-525-864A-1	Sequence 1, Appli
	157	15	1.0	3505	1	US-07-718-575-3	Sequence 3, Appli
	158	15	1.0	3505	1	US-08-481-206-3	Sequence 3, Appli
	159	15	1.0	3505	2	US-08-486-269A-3	Sequence 3, Appli
	160	15	1.0	3627	1	US-08-447-500-23	Sequence 23, Appl
	161	15	1.0	3627	1	US-08-454-097-23	Sequence 23, Appl
	162	15	1.0	3627	1	US-08-453-866-23	Sequence 23, Appl
	163	15	1.0	3627	3	US-08-185-359-23	Sequence 23, Appl
	164	15	1.0	3628	1	US-08-468-036-17	Sequence 17, Appl
	165	15	1.0	3628	2	US-08-376-843-17	Sequence 17, Appl
	166	15	1.0	3804	4	US-08-617-785-3	Sequence 3, Appli
	167	15	1.0	3825	4	US-09-208-742-3	Sequence 3, Appli
c	168	15	1.0	3937	3	US-08-586-165-8	Sequence 8, Appli
c	169	15	1.0	4500	1	US-08-308-872B-1	Sequence 1, Appli
c	170	15	1.0	4656	4	US-09-425-665-1	Sequence 1, Appli
	171	15	1.0	4673	1	US-07-638-431-1	Sequence 1, Appli
	172	15	1.0	4673	5	PCT-US92-00018-1	Sequence 1, Appli
	173	15	1.0	4975	2	US-08-249-687C-1	Sequence 1, Appli
	174	15	1.0	4989	2	US-08-666-392A-3	Sequence 3, Appli
	175	15	1.0	4989	2	US-08-625-819-1	Sequence 1, Appli
	176	15	1.0	4989	3	US-08-755-558-4	Sequence 4, Appli
	177	15	1.0	4989	3	US-08-746-559A-1	Sequence 1, Appli
	178	15	1.0	4993	3	US-08-746-559A-3	Sequence 3, Appli
	179	15	1.0	5051	2	US-08-239-276-2	Sequence 2, Appli
	180	15	1.0	5051	2	US-08-468-579B-2	Sequence 2, Appli
	181	15	1.0	5051	3	US-08-468-577B-2	Sequence 2, Appli
c	182	15	1.0	5157	2	US-08-474-169-7	Sequence 7, Appli
c	183	15	1.0	5405	3	US-09-282-996-1	Sequence 1, Appli
	184	15	1.0	5406	1	US-07-813-593-3	Sequence 3, Appli
	185	15	1.0	5406	1	US-07-977-451-5	Sequence 5, Appli
	186	15	1.0	5406	1	US-07-946-507-3	Sequence 3, Appli
	187	15	1.0	5406	1	US-08-252-517-5	Sequence 5, Appli
	188	15	1.0	5406	1	US-07-906-397A-5	Sequence 5, Appli
	189	15	1.0	5406	1	US-08-601-891-5	Sequence 5, Appli
	190	15	1.0	5406	2	US-09-021-324-5	Sequence 5, Appli
	191	15	1.0	5406	5	PCT-US92-02750-7	Sequence 7, Appli
	192	15	1.0	5406	5	PCT-US92-05401-5	Sequence 5, Appli
	193	15	1.0	5406	5	PCT-US92-09893-5	Sequence 5, Appli
c	194	15	1.0	5427	3	US-09-282-996-2	Sequence 2, Appli
	195	15	1.0	5470	2	US-08-443-861-1	Sequence 1, Appli
	196	15	1.0	5470	4	US-08-193-829B-1	Sequence 1, Appli
c	197	15	1.0	5559	1	US-08-287-442-3	Sequence 3, Appli
c	198	15	1.0	5559	1	US-08-459-701-3	Sequence 3, Appli
c	199	15	1.0	5559	1	US-08-460-298-3	Sequence 3, Appli
c	200	15	1.0	5559	1	US-08-459-174-3	Sequence 3, Appli
c	201	15	1.0	5559	5	PCT-US93-06300A-3	Sequence 3, Appli
c	202	15	1.0	5595	1	US-07-841-655-1	Sequence 1, Appli
c	203	15	1.0	5595	1	US-07-635-965C-1	Sequence 1, Appli
c	204	15	1.0	5620	2	US-08-793-170-21	Sequence 21, Appl
c	205	15	1.0	5620	3	US-08-892-873-21	Sequence 21, Appl
c	206	15	1.0	5620	4	US-09-334-765A-21	Sequence 21, Appl
	207	15	1.0	5648	3	US-09-371-008-1	Sequence 1, Appli
c	208	15	1.0	5789	3	US-08-862-431-32	Sequence 32, Appl
c	209	15	1.0	5791	3	US-08-862-431-31	Sequence 31, Appl
c	210	15	1.0	5793	3	US-08-862-431-29	Sequence 29, Appl
c	211	15	1.0	5793	3	US-08-862-431-30	Sequence 30, Appl
c	212	15	1.0	5818	2	US-08-536-559A-3	Sequence 3, Appli
c	213	15	1.0	5819	2	US-08-536-559A-2	Sequence 2, Appli
c	214	15	1.0	5819	3	US-08-862-431-27	Sequence 27, Appl
c	215	15	1.0	5819	3	US-08-862-431-28	Sequence 28, Appl
c	216	15	1.0	5938	2	US-08-536-559A-4	Sequence 4, Appli
c	217	15	1.0	6044	1	US-08-316-950-18	Sequence 18, Appl
c	218	15	1.0	6044	5	PCT-US95-12642-18	Sequence 18, Appl
	219	15	1.0	6047	1	US-08-316-950-12	Sequence 12, Appl
	220	15	1.0	6047	5	PCT-US95-12642-12	Sequence 12, Appl
c	221	15	1.0	6092	2	US-08-536-559A-1	Sequence 1, Appli
c	222	15	1.0	6092	3	US-08-862-431-26	Sequence 26, Appl
c	223	15	1.0	6171	1	US-08-316-950-17	Sequence 17, Appl
c	224	15	1.0	6171	5	PCT-US95-12642-17	Sequence 17, Appl
c	225	15	1.0	6525	1	US-08-493-092-3	Sequence 3, Appli
c	226	15	1.0	6525	1	US-08-508-836A-3	Sequence 3, Appli
	227	15	1.0	6831	2	US-08-609-049A-27	Sequence 27, Appl
	228	15	1.0	7070	1	US-08-619-554-3	Sequence 3, Appli
	229	15	1.0	7904	1	US-08-316-239B-1	Sequence 1, Appli
	230	15	1.0	7904	1	US-08-316-239B-2	Sequence 2, Appli
	231	15	1.0	7904	1	US-08-410-005-1	Sequence 1, Appli
	232	15	1.0	7904	3	US-08-929-140-1	Sequence 1, Appli
c	233	15	1.0	8220	2	US-08-568-459A-11	Sequence 11, Appl
c	234	15	1.0	8220	2	US-08-487-826B-11	Sequence 11, Appl
	235	15	1.0	9388	4	US-08-991-789A-141	Sequence 141, App
	236	15	1.0	10754	2	US-08-966-958-1	Sequence 1, Appli
	237	15	1.0	10754	2	US-09-215-817-1	Sequence 1, Appli
	238	15	1.0	10754	4	US-09-342-353-1	Sequence 1, Appli
c	239	15	1.0	10763	1	US-08-761-258-1	Sequence 1, Appli
c	240	15	1.0	10763	2	US-08-977-306-1	Sequence 1, Appli
c	241	15	1.0	11461	3	US-08-669-161A-29	Sequence 29, Appl
c	242	15	1.0	11601	2	US-08-222-617A-3	Sequence 3, Appli
c	243	15	1.0	11601	2	US-08-222-617A-24	Sequence 24, Appl
c	244	15	1.0	11616	1	US-08-196-259-2	Sequence 2, Appli
c	245	15	1.0	11785	1	US-08-038-768A-4	Sequence 4, Appli
c	246	15	1.0	11785	2	US-08-416-603-3	Sequence 3, Appli

c 247	15	1.0	11970	4	US-09-345-217-1	Sequence 1, Appli
248	15	1.0	13417	2	US-08-637-759B-37	Sequence 37, Appl
249	15	1.0	13417	3	US-08-871-355A-37	Sequence 37, Appl
c 250	15	1.0	14113	3	US-09-223-134-1	Sequence 1, Appli
c 251	15	1.0	14113	3	US-08-992-801-1	Sequence 1, Appli
c 252	15	1.0	14113	4	US-09-223-535-1	Sequence 1, Appli
253	15	1.0	14176	1	US-08-307-499-1	Sequence 1, Appli
c 254	15	1.0	14176	1	US-08-307-499-14	Sequence 14, Appl
255	15	1.0	14176	4	US-09-299-268-1	Sequence 1, Appli
c 256	15	1.0	14176	4	US-09-299-268-14	Sequence 14, Appl
257	15	1.0	14636	4	US-09-173-914-6	Sequence 6, Appli
258	15	1.0	14683	2	US-08-819-866-1	Sequence 1, Appli
259	15	1.0	14683	2	US-09-023-715-1	Sequence 1, Appli
c 260	15	1.0	17041	1	US-08-076-011-1	Sequence 1, Appli
261	15	1.0	18986	2	US-08-819-866-2	Sequence 2, Appli
262	15	1.0	18986	2	US-09-023-715-2	Sequence 2, Appli
c 263	15	1.0	19124	2	US-08-487-826B-13	Sequence 13, Appl
c 264	15	1.0	24979	2	US-08-147-777-3	Sequence 3, Appli
c 265	15	1.0	24979	3	US-08-452-872-3	Sequence 3, Appli
c 266	15	1.0	24979	5	PCT-US93-03985-3	Sequence 3, Appli
c 267	15	1.0	35060	3	US-08-814-095-7	Sequence 7, Appli
268	15	1.0	36519	3	US-08-923-137-2	Sequence 2, Appli
269	15	1.0	43676	3	US-09-356-952-12	Sequence 12, Appl
270	15	1.0	43795	3	US-08-742-185-101	Sequence 101, App
c 271	15	1.0	46899	1	US-08-471-119A-1	Sequence 1, Appli
272	15	1.0	51952	3	US-08-947-823-1	Sequence 1, Appli
c 273	15	1.0	53526	3	US-08-658-136-2	Sequence 2, Appli
c 274	15	1.0	53577	3	US-08-658-136-1	Sequence 1, Appli
c 275	15	1.0	87350	3	US-08-781-891-79	Sequence 79, Appl
c 276	15	1.0	152331	3	US-09-128-155-16	Sequence 16, Appl
c 277	15	1.0	176373	3	US-09-128-155-17	Sequence 17, Appl
c 278	14	0.9	15	1	US-08-319-492B-17	Sequence 17, Appl
c 279	14	0.9	17	1	US-08-233-608-8	Sequence 8, Appli
c 280	14	0.9	17	1	US-08-887-480-8	Sequence 8, Appli
c 281	14	0.9	17	2	US-08-722-187-8	Sequence 8, Appli
c 282	14	0.9	17	5	PCT-US95-04712-8	Sequence 8, Appli
c 283	14	0.9	18	3	US-08-392-794A-11	Sequence 11, Appl
284	14	0.9	21	1	US-08-185-301-4	Sequence 4, Appli
c 285	14	0.9	21	3	US-09-032-894-114	Sequence 114, App
286	14	0.9	21	3	US-09-032-894-116	Sequence 116, App
c 287	14	0.9	21	4	US-09-031-626-114	Sequence 114, App
288	14	0.9	21	4	US-09-031-626-116	Sequence 116, App
c 289	14	0.9	23	1	US-08-233-608-9	Sequence 9, Appli
c 290	14	0.9	23	1	US-08-887-480-9	Sequence 9, Appli
c 291	14	0.9	23	2	US-08-722-187-9	Sequence 9, Appli
c 292	14	0.9	23	5	PCT-US95-04712-9	Sequence 9, Appli
c 293	14	0.9	28	1	US-08-647-584-51	Sequence 51, Appl
c 294	14	0.9	29	1	US-08-647-584-52	Sequence 52, Appl
c 295	14	0.9	31	3	US-09-032-894-115	Sequence 115, App
296	14	0.9	31	3	US-09-032-894-117	Sequence 117, App
c 297	14	0.9	31	4	US-09-031-626-115	Sequence 115, App
298	14	0.9	31	4	US-09-031-626-117	Sequence 117, App
299	14	0.9	33	1	US-08-402-006-2	Sequence 2, Appli
c 300	14	0.9	36	1	US-08-334-847-836	Sequence 836,.App
c 301	14	0.9	39	1	US-08-466-852-1	Sequence 1, Appli
302	14	0.9	47	1	US-08-466-852-2	Sequence 2, Appli
303	14	0.9	50	1	US-08-171-389-161	Sequence 161, App
304	14	0.9	50	1	US-08-123-936-161	Sequence 161, App
c 305	14	0.9	50	1	US-08-374-641-32	Sequence 32, Appl
306	14	0.9	50	2	US-08-475-228A-161	Sequence 161, App
307	14	0.9	50	3	US-08-482-080A-161	Sequence 161, App
308	14	0.9	50	5	PCT-US93-12388-161	Sequence 161, App
c 309	14	0.9	87	5	PCT-US95-05600-6	Sequence 6, Appli
c 310	14	0.9	87	5	PCT-US95-05600-11	Sequence 11, Appl
c 311	14	0.9	106	2	US-08-852-807-15	Sequence 15, Appl
312	14	0.9	152	1	US-08-136-277-5	Sequence 5, Appli
313	14	0.9	152	2	US-08-479-403-5	Sequence 5, Appli
314	14	0.9	152	3	US-08-835-734-5	Sequence 5, Appli
c 315	14	0.9	167	1	US-08-392-678-24	Sequence 24, Appl
c 316	14	0.9	167	1	US-08-457-304A-24	Sequence 24, Appl
c 317	14	0.9	167	1	US-08-456-701A-24	Sequence 24, Appl
318	14	0.9	181	1	US-08-480-552-3	Sequence 3, Appli
319	14	0.9	181	1	US-08-039-385-3	Sequence 3, Appli
320	14	0.9	181	3	US-08-929-208-3	Sequence 3, Appli
321	14	0.9	181	4	US-09-158-469-3	Sequence 3, Appli
322	14	0.9	181	5	PCT-US91-07492-3	Sequence 3, Appli
323	14	0.9	206	1	US-08-480-552-7	Sequence 7, Appli
324	14	0.9	206	1	US-08-039-385-7	Sequence 7, Appli
325	14	0.9	206	3	US-08-929-208-7	Sequence 7, Appli
326	14	0.9	206	4	US-09-158-469-7	Sequence 7, Appli
327	14	0.9	206	5	PCT-US91-07492-7	Sequence 7, Appli
c 328	14	0.9	239	4	US-09-060-756-376	Sequence 376, App
c 329	14	0.9	254	1	US-08-171-718-117	Sequence 117, App
c 330	14	0.9	254	3	US-08-478-087-117	Sequence 117, App
331	14	0.9	274	4	US-08-905-223-185	Sequence 185, App
332	14	0.9	276	6	5171841-3	Patent No. 5171841
333	14	0.9	278	2	US-08-637-759B-33	Sequence 33, Appl
334	14	0.9	278	3	US-08-871-355A-33	Sequence 33, Appl
335	14	0.9	278	4	US-09-060-756-210	Sequence 210, App
336	14	0.9	288	6	5171841-2	Patent No. 5171841
c 337	14	0.9	294	4	US-08-991-789A-20	Sequence 20, Appl
338	14	0.9	300	2	US-08-637-759B-9	Sequence 9, Appli
339	14	0.9	300	3	US-08-871-355A-9	Sequence 9, Appli
340	14	0.9	316	2	US-09-032-684-10	Sequence 10, Appl
341	14	0.9	321	3	US-08-483-749A-11	Sequence 11, Appl
342	14	0.9	324	4	US-08-905-223-264	Sequence 264, App
c 343	14	0.9	340	3	US-08-441-971-14	Sequence 14, Appl
c 344	14	0.9	340	4	US-08-221-653-14	Sequence 14, Appl
c 345	14	0.9	340	4	US-08-442-144A-14	Sequence 14, Appl
c 346	14	0.9	346	4	US-09-060-756-385	Sequence 385, App
347	14	0.9	348	1	US-08-216-276A-26	Sequence 26, Appl
348	14	0.9	384	2	US-08-518-862C-3	Sequence 3, Appli
c 349	14	0.9	391	1	US-08-592-126-83	Sequence 83, Appl
350	14	0.9	410	4	US-09-276-531-88	Sequence 88, Appl
c 351	14	0.9	420	4	US-09-060-756-417	Sequence 417, App
352	14	0.9	442	1	US-08-229-781-29	Sequence 29, Appl
353	14	0.9	442	1	US-08-630-918-29	Sequence 29, Appl
354	14	0.9	443	4	US-08-998-416-195	Sequence 195, App
355	14	0.9	444	2	US-09-070-060-11	Sequence 11, Appl
356	14	0.9	444	3	US-09-357-746-11	Sequence 11, Appl
357	14	0.9	446	2	US-08-332-766A-26	Sequence 26, Appl
c 358	14	0.9	456	1	US-07-961-724C-15	Sequence 15, Appl
c 359	14	0.9	461	3	US-09-034-810-3	Sequence 3, Appli
c 360	14	0.9	461	3	US-08-685-239-3	Sequence 3, Appli
361	14	0.9	461	4	US-09-257-584-7	Sequence 7, Appli
362	14	0.9	470	3	US-09-032-894-97	Sequence 97, Appl
363	14	0.9	470	4	US-09-031-626-97	Sequence 97, Appl
364	14	0.9	474	5	PCT-US94-01149-33	Sequence 33, Appl
365	14	0.9	481	1	US-08-136-277-19	Sequence 19, Appl
366	14	0.9	481	2	US-08-479-403-19	Sequence 19, Appl
367	14	0.9	481	3	US-08-835-734-19	Sequence 19, Appl
c 368	14	0.9	495	1	US-08-249-671A-6	Sequence 6, Appli
c 369	14	0.9	498	2	US-08-631-328-54	Sequence 54, Appl
370	14	0.9	528	3	US-08-714-918-24	Sequence 24, Appl
371	14	0.9	528	4	US-09-265-315-24	Sequence 24, Appl
372	14	0.9	528	4	US-09-265-315-24	Sequence 24, Appl
373	14	0.9	528	4	US-09-266-417-24	Sequence 24, Appl
c 374	14	0.9	540	6	5310729-1	Patent No. 5310729
375	14	0.9	558	4	US-08-998-416-906	Sequence 906, App
376	14	0.9	563	1	US-08-887-480-86	Sequence 86, Appl
377	14	0.9	563	2	US-08-722-187-86	Sequence 86, Appl
378	14	0.9	563	5	PCT-US95-04712-86	Sequence 86, Appl
379	14	0.9	582	1	US-08-464-342-5	Sequence 5, Appli
380	14	0.9	582	2	US-08-464-604A-5	Sequence 5, Appli
381	14	0.9	582	2	US-08-875-272-5	Sequence 5, Appli
382	14	0.9	582	2	US-08-903-396-5	Sequence 5, Appli
383	14	0.9	583	1	US-08-233-608-2	Sequence 2, Appli
384	14	0.9	583	1	US-08-887-480-2	Sequence 2, Appli
385	14	0.9	583	2	US-08-722-187-2	Sequence 2, Appli
386	14	0.9	583	5	PCT-US95-04712-2	Sequence 2, Appli
c 387	14	0.9	584	1	US-08-181-271A-17	Sequence 17, Appl
c 388	14	0.9	584	1	US-08-449-315-17	Sequence 17, Appl
c 389	14	0.9	584	1	US-08-444-803-17	Sequence 17, Appl
c 390	14	0.9	584	1	US-08-449-043-17	Sequence 17, Appl
c 391	14	0.9	584	1	US-08-456-265A-15	Sequence 15, Appl
c 392	14	0.9	584	1	US-08-455-416-17	Sequence 17, Appl

c 393	14	0.9	584	1	US-08-455-244-17	Sequence 17, Appl
c 394	14	0.9	584	1	US-08-454-876-17	Sequence 17, Appl
c 395	14	0.9	584	2	US-08-457-364-17	Sequence 17, Appl
c 396	14	0.9	584	2	US-08-456-262-17	Sequence 17, Appl
c 397	14	0.9	584	2	US-08-456-240-17	Sequence 17, Appl
c 398	14	0.9	584	2	US-08-455-736-17	Sequence 17, Appl
c 399	14	0.9	584	2	US-08-971-217-15	Sequence 15, Appl
c 400	14	0.9	591	2	US-08-704-473-1	Sequence 1, Appli
401	14	0.9	596	2	US-08-105-989-3	Sequence 3, Appli
402	14	0.9	596	3	US-09-138-922-3	Sequence 3, Appli
403	14	0.9	618	1	US-08-155-171B-5	Sequence 5, Appli
404	14	0.9	618	2	US-08-435-998-5	Sequence 5, Appli
405	14	0.9	630	5	PCT-US94-01149-21	Sequence 21, Appl
406	14	0.9	631	4	US-08-642-274D-33	Sequence 33, Appl
407	14	0.9	634	1	US-08-482-069-1	Sequence 1, Appli
408	14	0.9	634	1	US-08-486-410-1	Sequence 1, Appli
409	14	0.9	634	1	US-08-263-935-1	Sequence 1, Appli
c 410	14	0.9	662	1	US-07-807-043B-10	Sequence 10, Appl
c 411	14	0.9	662	1	US-08-299-849B-10	Sequence 10, Appl
c 412	14	0.9	662	2	US-08-142-368A-10	Sequence 10, Appl
c 413	14	0.9	662	3	US-08-967-727-10	Sequence 10, Appl
c 414	14	0.9	662	4	US-08-037-230D-10	Sequence 10, Appl
415	14	0.9	670	5	PCT-US94-01149-5	Sequence 5, Appli
416	14	0.9	675	5	PCT-US94-14498A-1	Sequence 1, Appli
417	14	0.9	687	1	US-07-916-901-7	Sequence 7, Appli
c 418	14	0.9	699	3	US-09-361-434-9	Sequence 9, Appli
419	14	0.9	699	3	US-09-361-434-10	Sequence 10, Appl
c 420	14	0.9	701	4	US-08-998-416-833	Sequence 833, App
421	14	0.9	705	1	US-08-506-404D-2	Sequence 2, Appli
422	14	0.9	705	3	US-09-035-754-2	Sequence 2, Appli
423	14	0.9	706	4	US-08-998-416-873	Sequence 873, App
424	14	0.9	717	5	PCT-US94-01149-23	Sequence 23, Appl
c 425	14	0.9	718	4	US-08-998-416-682	Sequence 682, App
c 426	14	0.9	720	6	5310729-3	Patent No. 5310729
c 427	14	0.9	724	3	US-08-961-083-45	Sequence 45, Appl
428	14	0.9	729	5	PCT-US94-01149-17	Sequence 17, Appl
429	14	0.9	734	3	US-09-014-583-1	Sequence 1, Appli
430	14	0.9	735	3	US-08-950-720A-5	Sequence 5, Appli
431	14	0.9	747	2	US-08-679-765-2	Sequence 2, Appli
432	14	0.9	747	2	US-09-196-525-2	Sequence 2, Appli
433	14	0.9	747	4	US-09-318-317-2	Sequence 2, Appli
c 434	14	0.9	768	1	US-08-236-918A-5	Sequence 5, Appli
c 435	14	0.9	775	3	US-09-361-434-6	Sequence 6, Appli
436	14	0.9	775	3	US-09-361-434-8	Sequence 8, Appli
437	14	0.9	781	4	US-08-998-416-214	Sequence 214, App
438	14	0.9	790	1	US-08-393-985-22	Sequence 22, Appl
c 439	14	0.9	792	1	US-08-578-709-10	Sequence 10, Appl
440	14	0.9	798	2	US-08-742-009-2	Sequence 2, Appli
441	14	0.9	798	3	US-09-172-110-2	Sequence 2, Appli
442	14	0.9	805	1	US-08-118-469A-6	Sequence 6, Appli
443	14	0.9	805	1	US-08-909-119-6	Sequence 6, Appli
444	14	0.9	810	5	PCT-US94-01149-19	Sequence 19, Appl
c 445	14	0.9	816	3	US-09-079-839-2	Sequence 2, Appli
c 446	14	0.9	841	4	US-08-998-416-318	Sequence 318, App
c 447	14	0.9	858	2	US-08-964-631-3	Sequence 3, Appli
c 448	14	0.9	858	4	US-09-210-000-3	Sequence 3, Appli
c 449	14	0.9	869	1	US-08-249-671A-7	Sequence 7, Appli
c 450	14	0.9	875	2	US-08-867-676-2	Sequence 2, Appli
c 451	14	0.9	878	2	US-08-446-345-33	Sequence 33, Appl
c 452	14	0.9	879	1	US-08-249-671A-10	Sequence 10, Appl
453	14	0.9	886	2	US-08-827-190-9	Sequence 9, Appli
454	14	0.9	900	2	US-08-518-862C-5	Sequence 5, Appli
455	14	0.9	912	5	PCT-US94-01149-31	Sequence 31, Appl
456	14	0.9	924	5	PCT-US94-01149-15	Sequence 15, Appl
c 457	14	0.9	966	3	US-09-361-434-19	Sequence 19, Appl
458	14	0.9	966	3	US-09-361-434-20	Sequence 20, Appl
c 459	14	0.9	989	3	US-09-361-434-21	Sequence 21, Appl
460	14	0.9	989	3	US-09-361-434-23	Sequence 23, Appl
c 461	14	0.9	990	3	US-08-964-308-12	Sequence 12, Appl
c 462	14	0.9	990	3	US-08-964-308-13	Sequence 13, Appl
c 463	14	0.9	990	3	US-08-964-313-12	Sequence 12, Appl
c 464	14	0.9	990	3	US-08-964-313-13	Sequence 13, Appl
465	14	0.9	999	1	US-08-469-649-1	Sequence 1, Appli
c 466	14	0.9	1013	4	US-09-084-120-17	Sequence 17, Appl
467	14	0.9	1014	5	PCT-US92-10421-1	Sequence 1, Appli
468	14	0.9	1017	4	US-08-793-634B-6	Sequence 6, Appli
c 469	14	0.9	1018	3	US-08-896-095-2	Sequence 2, Appli
c 470	14	0.9	1045	1	US-08-317-880-3	Sequence 3, Appli
c 471	14	0.9	1045	2	US-08-782-396-3	Sequence 3, Appli
c 472	14	0.9	1065	3	US-08-591-685-6	Sequence 6, Appli
c 473	14	0.9	1077	1	US-08-409-199-4	Sequence 4, Appli
c 474	14	0.9	1079	1	US-07-809-457A-1	Sequence 1, Appli
c 475	14	0.9	1079	1	US-08-553-943-1	Sequence 1, Appli
c 476	14	0.9	1079	5	PCT-US91-09437-1	Sequence 1, Appli
c 477	14	0.9	1080	2	US-09-092-770-5	Sequence 5, Appli
c 478	14	0.9	1080	4	US-09-222-851-5	Sequence 5, Appli
c 479	14	0.9	1091	1	US-08-452-531-1	Sequence 1, Appli
c 480	14	0.9	1091	1	US-08-460-746A-1	Sequence 1, Appli
c 481	14	0.9	1091	1	US-08-460-555-1	Sequence 1, Appli
c 482	14	0.9	1091	3	US-08-460-066-1	Sequence 1, Appli
c 483	14	0.9	1106	3	US-09-361-434-16	Sequence 16, Appl
484	14	0.9	1106	3	US-09-361-434-18	Sequence 18, Appl
c 485	14	0.9	1108	2	US-08-852-807-2	Sequence 2, Appli
c 486	14	0.9	1116	2	US-08-805-965-4	Sequence 4, Appli
c 487	14	0.9	1122	1	US-08-599-171A-18	Sequence 18, Appl
c 488	14	0.9	1122	2	US-08-646-590B-18	Sequence 18, Appl
c 489	14	0.9	1122	3	US-09-069-226-18	Sequence 18, Appl
c 490	14	0.9	1122	4	US-09-412-184-18	Sequence 18, Appl
c 491	14	0.9	1167	1	US-07-960-985-1	Sequence 1, Appli
c 492	14	0.9	1167	2	US-08-496-671-1	Sequence 1, Appli
493	14	0.9	1196	3	US-08-891-789B-1	Sequence 1, Appli
c 494	14	0.9	1197	2	US-08-774-104A-1	Sequence 1, Appli
495	14	0.9	1200	1	US-08-356-397-1	Sequence 1, Appli
496	14	0.9	1201	1	US-08-207-904-9	Sequence 9, Appli
c 497	14	0.9	1208	3	US-08-860-519-7	Sequence 7, Appli
498	14	0.9	1212	2	US-08-802-322-4	Sequence 4, Appli
c 499	14	0.9	1215	2	US-09-092-770-1	Sequence 1, Appli
c 500	14	0.9	1215	4	US-09-222-851-1	Sequence 1, Appli
501	14	0.9	1240	1	US-08-388-756-1	Sequence 1, Appli
502	14	0.9	1240	2	US-08-748-428-1	Sequence 1, Appli
503	14	0.9	1244	5	PCT-US96-02331-12	Sequence 12, Appl
c 504	14	0.9	1257	2	US-08-964-631-1	Sequence 1, Appli
c 505	14	0.9	1257	4	US-09-210-000-1	Sequence 1, Appli
506	14	0.9	1263	4	US-09-239-303-8	Sequence 8, Appli
507	14	0.9	1270	1	US-08-915-003-8	Sequence 8, Appli
508	14	0.9	1270	2	US-08-642-247-8	Sequence 8, Appli
509	14	0.9	1273	2	US-08-474-379C-25	Sequence 25, Appl
510	14	0.9	1273	3	US-09-146-249A-25	Sequence 25, Appl
511	14	0.9	1273	3	US-08-206-188B-25	Sequence 25, Appl
512	14	0.9	1276	1	US-07-688-352C-25	Sequence 25, Appl
513	14	0.9	1276	5	PCT-US91-02714-24	Sequence 24, Appl
c 514	14	0.9	1287	2	US-08-846-705-5	Sequence 5, Appli
515	14	0.9	1294	1	US-07-662-005A-3	Sequence 3, Appli
c 516	14	0.9	1302	1	US-08-719-331-7	Sequence 7, Appli
c 517	14	0.9	1302	3	US-09-100-391-11	Sequence 11, Appl
518	14	0.9	1309	2	US-08-933-750C-63	Sequence 63, Appl
519	14	0.9	1309	3	US-09-234-613-63	Sequence 63, Appl
c 520	14	0.9	1311	3	US-09-100-391-3	Sequence 3, Appli
c 521	14	0.9	1312	2	US-08-841-349-6	Sequence 6, Appli
522	14	0.9	1314	2	US-08-440-845D-9	Sequence 9, Appli
523	14	0.9	1314	3	US-08-868-458-9	Sequence 9, Appli
c 524	14	0.9	1323	1	US-08-307-499-36	Sequence 36, Appl
c 525	14	0.9	1323	4	US-09-299-268-36	Sequence 36, Appl
c 526	14	0.9	1333	3	US-08-889-425-3	Sequence 3, Appli
c 527	14	0.9	1362	6	5290690-8	Patent No. 5290690
528	14	0.9	1380	2	US-08-748-947A-1	Sequence 1, Appli
529	14	0.9	1384	2	US-08-860-577-1	Sequence 1, Appli
530	14	0.9	1392	4	US-09-163-444-1	Sequence 1, Appli
531	14	0.9	1419	2	US-08-194-981E-4	Sequence 4, Appli
532	14	0.9	1427	3	US-09-173-581-10	Sequence 10, Appl
533	14	0.9	1428	2	US-08-960-022-7	Sequence 7, Appli
534	14	0.9	1432	5	PCT-US94-12912-4	Sequence 4, Appli
c 535	14	0.9	1440	2	US-08-743-637B-174	Sequence 174, App
c 536	14	0.9	1440	3	US-08-526-840B-174	Sequence 174, App
537	14	0.9	1443	4	US-09-078-173A-11	Sequence 11, Appl
538	14	0.9	1444	3	US-09-416-213-1	Sequence 1, Appli

	539	14	0.9	1456	3	US-09-308-406-1	Sequence 1, Appli
	540	14	0.9	1462	4	US-08-975-215-3	Sequence 3, Appli
c	541	14	0.9	1482	1	US-08-330-121B-1	Sequence 1, Appli
c	542	14	0.9	1482	5	PCT-US95-13820-1	Sequence 1, Appli
	543	14	0.9	1500	1	US-08-476-008-6	Sequence 6, Appli
	544	14	0.9	1500	1	US-08-306-063-6	Sequence 6, Appli
	545	14	0.9	1500	1	US-08-833-485-6	Sequence 6, Appli
	546	14	0.9	1500	4	US-09-137-440-6	Sequence 6, Appli
	547	14	0.9	1500	5	PCT-US91-06148A-6	Sequence 6, Appli
	548	14	0.9	1505	2	US-08-909-965C-13	Sequence 13, Appl
c	549	14	0.9	1529	2	US-08-477-451-17	Sequence 17, Appl
	550	14	0.9	1529	2	US-08-477-451-21	Sequence 21, Appl
c	551	14	0.9	1541	6	5183884-1	Patent No. 5183884
c	552	14	0.9	1542	1	US-07-978-895-1	Sequence 1, Appli
c	553	14	0.9	1542	1	US-08-473-119-1	Sequence 1, Appli
c	554	14	0.9	1542	2	US-08-475-352-1	Sequence 1, Appli
	555	14	0.9	1558	1	US-08-416-870C-9	Sequence 9, Appli
c	556	14	0.9	1570	3	US-08-617-860B-20	Sequence 20, Appl
c	557	14	0.9	1571	4	US-09-276-531-49	Sequence 49, Appl
c	558	14	0.9	1584	1	US-08-247-908A-10	Sequence 10, Appl
c	559	14	0.9	1584	1	US-08-453-942-10	Sequence 10, Appl
c	560	14	0.9	1584	2	US-08-926-885A-10	Sequence 10, Appl
c	561	14	0.9	1584	5	PCT-US94-05290-10	Sequence 10, Appl
	562	14	0.9	1591	2	US-08-194-981E-3	Sequence 3, Appli
	563	14	0.9	1608	4	US-09-029-348-19	Sequence 19, Appl
c	564	14	0.9	1619	1	US-08-208-007A-1	Sequence 1, Appli
c	565	14	0.9	1619	5	PCT-US94-04781-1	Sequence 1, Appli
c	566	14	0.9	1641	3	US-08-860-519-1	Sequence 1, Appli
	567	14	0.9	1665	4	US-08-591-468-6	Sequence 6, Appli
	568	14	0.9	1665	5	PCT-US94-06430-6	Sequence 6, Appli
	569	14	0.9	1667	1	US-08-485-284A-1	Sequence 1, Appli
c	570	14	0.9	1669	3	US-08-964-308-3	Sequence 3, Appli
c	571	14	0.9	1669	3	US-08-964-313-3	Sequence 3, Appli
c	572	14	0.9	1672	1	US-07-721-761A-27	Sequence 27, Appl
c	573	14	0.9	1672	1	US-07-978-687-27	Sequence 27, Appl
c	574	14	0.9	1672	5	PCT-US91-05801-27	Sequence 27, Appl
	575	14	0.9	1673	1	US-08-476-008-4	Sequence 4, Appli
	576	14	0.9	1673	1	US-08-306-063-4	Sequence 4, Appli
	577	14	0.9	1673	1	US-08-833-485-4	Sequence 4, Appli
	578	14	0.9	1673	4	US-09-137-440-4	Sequence 4, Appli
	579	14	0.9	1673	5	PCT-US91-06148A-4	Sequence 4, Appli
	580	14	0.9	1674	3	US-09-059-369-17	Sequence 17, Appl
c	581	14	0.9	1700	3	US-08-860-519-5	Sequence 5, Appli
c	582	14	0.9	1700	3	US-08-860-519-6	Sequence 6, Appli
	583	14	0.9	1707	3	US-08-976-852A-1	Sequence 1, Appli
c	584	14	0.9	1728	1	US-08-403-866-12	Sequence 12, Appl
c	585	14	0.9	1734	1	US-08-551-437-4	Sequence 4, Appli
c	586	14	0.9	1734	3	US-09-004-225-4	Sequence 4, Appli
c	587	14	0.9	1734	4	US-09-084-346-4	Sequence 4, Appli
	588	14	0.9	1735	4	US-09-163-444-3	Sequence 3, Appli
c	589	14	0.9	1784	6	5180810-2	Patent No. 5180810
	590	14	0.9	1797	1	US-08-181-556-1	Sequence 1, Appli
	591	14	0.9	1797	1	US-08-181-556-3	Sequence 3, Appli
	592	14	0.9	1797	1	US-08-323-084A-9	Sequence 9, Appli
	593	14	0.9	1797	1	US-08-674-008-9	Sequence 9, Appli
c	594	14	0.9	1799	4	US-09-329-633A-1	Sequence 1, Appli
	595	14	0.9	1810	4	US-08-591-468-10	Sequence 10, Appl
	596	14	0.9	1810	5	PCT-US94-06430-10	Sequence 10, Appl
	597	14	0.9	1839	1	US-08-075-193-1	Sequence 1, Appli
	598	14	0.9	1839	2	US-08-564-090A-1	Sequence 1, Appli
	599	14	0.9	1839	5	PCT-US94-06698-1	Sequence 1, Appli
	600	14	0.9	1841	2	US-08-820-521-1	Sequence 1, Appli
	601	14	0.9	1841	4	US-09-248-715-1	Sequence 1, Appli
c	602	14	0.9	1845	1	US-07-721-761A-28	Sequence 28, Appl
c	603	14	0.9	1845	1	US-07-978-687-28	Sequence 28, Appl
c	604	14	0.9	1845	5	PCT-US91-05801-28	Sequence 28, Appl
c	605	14	0.9	1846	2	US-08-663-566A-16	Sequence 16, Appl
c	606	14	0.9	1846	2	US-08-023-610-16	Sequence 16, Appl
c	607	14	0.9	1846	2	US-08-288-065A-16	Sequence 16, Appl
c	608	14	0.9	1846	2	US-08-362-240A-16	Sequence 16, Appl
c	609	14	0.9	1846	5	PCT-US95-10245-16	Sequence 16, Appl
	610	14	0.9	1852	1	US-08-201-118-10	Sequence 10, Appl
	611	14	0.9	1852	2	US-08-238-821B-10	Sequence 10, Appl
	612	14	0.9	1852	5	PCT-US95-05744-10	Sequence 10, Appl
	613	14	0.9	1854	1	US-08-201-118-4	Sequence 4, Appli
	614	14	0.9	1854	2	US-08-238-821B-4	Sequence 4, Appli
	615	14	0.9	1854	5	PCT-US95-05744-4	Sequence 4, Appli
c	616	14	0.9	1870	3	US-09-295-186-9	Sequence 9, Appli
	617	14	0.9	1871	4	US-09-104-068-3	Sequence 3, Appli
c	618	14	0.9	1875	5	PCT-US95-10166-1	Sequence 1, Appli
	619	14	0.9	1884	3	US-08-875-223-7	Sequence 7, Appli
c	620	14	0.9	1893	2	US-08-596-319-1	Sequence 1, Appli
c	621	14	0.9	1922	2	US-08-372-652-9	Sequence 9, Appli
c	622	14	0.9	1922	5	PCT-US95-16311-9	Sequence 9, Appli
	623	14	0.9	1947	1	US-08-299-849B-19	Sequence 19, Appl
	624	14	0.9	1947	2	US-08-142-368A-19	Sequence 19, Appl
	625	14	0.9	1947	3	US-08-967-727-19	Sequence 19, Appl
	626	14	0.9	1947	4	US-08-037-230D-19	Sequence 19, Appl
	627	14	0.9	1952	1	US-08-333-358-1	Sequence 1, Appli
	628	14	0.9	1952	1	US-08-463-694-1	Sequence 1, Appli
	629	14	0.9	1952	1	US-08-694-501-1	Sequence 1, Appli
	630	14	0.9	1960	1	US-08-415-823-3	Sequence 3, Appli
c	631	14	0.9	1960	2	US-08-595-937A-1	Sequence 1, Appli
	632	14	0.9	1960	2	US-09-086-662-3	Sequence 3, Appli
c	633	14	0.9	1960	3	US-09-165-240-4	Sequence 4, Appli
	634	14	0.9	1965	2	US-08-809-763-5	Sequence 5, Appli
	635	14	0.9	1965	3	US-08-956-253-5	Sequence 5, Appli
c	636	14	0.9	1977	1	US-08-578-709-14	Sequence 14, Appl
c	637	14	0.9	1985	1	US-07-792-865D-2	Sequence 2, Appli
	638	14	0.9	2004	2	US-08-691-814B-1	Sequence 1, Appli
c	639	14	0.9	2016	1	US-08-294-872-1	Sequence 1, Appli
c	640	14	0.9	2016	5	PCT-US95-09823-1	Sequence 1, Appli
c	641	14	0.9	2048	1	US-08-416-870C-7	Sequence 7, Appli
	642	14	0.9	2051	2	US-08-672-814D-1	Sequence 1, Appli
	643	14	0.9	2051	4	US-09-333-696-1	Sequence 1, Appli
c	644	14	0.9	2080	1	US-08-179-738-9	Sequence 9, Appli
c	645	14	0.9	2080	2	US-08-628-145-9	Sequence 9, Appli
	646	14	0.9	2090	3	US-09-059-369-1	Sequence 1, Appli
	647	14	0.9	2097	1	US-08-393-985-1	Sequence 1, Appli
c	648	14	0.9	2100	1	US-08-485-718-10	Sequence 10, Appl
c	649	14	0.9	2100	2	US-08-484-530-56	Sequence 56, Appl
c	650	14	0.9	2100	2	US-08-827-618A-56	Sequence 56, Appl
c	651	14	0.9	2100	3	US-08-483-952A-56	Sequence 56, Appl
	652	14	0.9	2100	4	US-09-104-068-1	Sequence 1, Appli
c	653	14	0.9	2100	6	5475086-5	Patent No. 5475086
c	654	14	0.9	2122	1	US-08-485-449-1	Sequence 1, Appli
	655	14	0.9	2136	1	US-08-321-587-1	Sequence 1, Appli
	656	14	0.9	2204	3	US-09-257-799-47	Sequence 47, Appl
	657	14	0.9	2204	3	US-08-920-919A-47	Sequence 47, Appl
c	658	14	0.9	2231	1	US-08-403-866-14	Sequence 14, Appl
c	659	14	0.9	2235	4	US-09-153-804-2	Sequence 2, Appli
c	660	14	0.9	2249	2	US-08-494-624-1	Sequence 1, Appli
c	661	14	0.9	2347	5	PCT-US96-03965-1	Sequence 1, Appli
	662	14	0.9	2371	2	US-09-020-466-1	Sequence 1, Appli
	663	14	0.9	2371	3	US-09-192-659-1	Sequence 1, Appli
	664	14	0.9	2500	1	US-08-550-715-10	Sequence 10, Appl
c	665	14	0.9	2529	2	US-08-867-129-1	Sequence 1, Appli
c	666	14	0.9	2643	3	US-09-100-391-9	Sequence 9, Appli
c	667	14	0.9	2653	1	US-08-295-670-5	Sequence 5, Appli
c	668	14	0.9	2653	1	US-08-633-485-5	Sequence 5, Appli
c	669	14	0.9	2676	3	US-09-212-971-11	Sequence 11, Appl
c	670	14	0.9	2676	3	US-08-800-929A-11	Sequence 11, Appl
	671	14	0.9	2682	4	US-09-020-465-1	Sequence 1, Appli
c	672	14	0.9	2693	1	US-08-295-670-7	Sequence 7, Appli
c	673	14	0.9	2693	1	US-08-633-485-7	Sequence 7, Appli
c	674	14	0.9	2746	2	US-09-016-000-12	Sequence 12, Appl
	675	14	0.9	2809	1	US-08-484-105-3	Sequence 3, Appli
	676	14	0.9	2809	1	US-08-484-106-3	Sequence 3, Appli
c	677	14	0.9	2816	1	US-08-785-241-1	Sequence 1, Appli
c	678	14	0.9	2905	4	US-08-942-686-1	Sequence 1, Appli
	679	14	0.9	2919	3	US-08-434-000A-3	Sequence 3, Appli
c	680	14	0.9	2923	1	US-08-377-292-6	Sequence 6, Appli
c	681	14	0.9	2923	2	US-07-989-847-7	Sequence 7, Appli
c	682	14	0.9	2923	4	US-08-469-411-7	Sequence 7, Appli
c	683	14	0.9	2923	6	5187076-5	Patent No. 5187076
c	684	14	0.9	2934	3	US-09-149-934-2	Sequence 2, Appli

c 685	14	0.9	2937	6	5208144-7	Patent No. 5208144
686	14	0.9	2965	2	US-08-460-570-1	Sequence 1, Appli
687	14	0.9	2965	2	US-08-460-570-2	Sequence 2, Appli
688	14	0.9	2965	3	US-08-286-870A-1	Sequence 1, Appli
689	14	0.9	2965	3	US-08-286-870A-2	Sequence 2, Appli
690	14	0.9	3000	1	US-08-393-985-3	Sequence 3, Appli
691	14	0.9	3016	1	US-07-805-123C-1	Sequence 1, Appli
692	14	0.9	3016	1	US-08-033-081B-1	Sequence 1, Appli
693	14	0.9	3027	2	US-09-132-619-9	Sequence 9, Appli
694	14	0.9	3027	3	US-09-282-803B-9	Sequence 9, Appli
695	14	0.9	3039	2	US-08-680-326-27	Sequence 27, Appl
696	14	0.9	3095	3	US-08-434-000A-7	Sequence 7, Appli
697	14	0.9	3113	2	US-08-993-228-20	Sequence 20, Appl
c 698	14	0.9	3210	3	US-08-613-009A-6	Sequence 6, Appli
699	14	0.9	3230	1	US-08-216-276A-18	Sequence 18, Appl
700	14	0.9	3230	1	US-08-219-262B-11	Sequence 11, Appl
701	14	0.9	3230	3	US-09-031-655-11	Sequence 11, Appl
c 702	14	0.9	3230	3	US-09-280-799-78	Sequence 78, Appl
703	14	0.9	3230	5	PCT-US91-03056-16	Sequence 16, Appl
c 704	14	0.9	3230	6	5324640-1	Patent No. 5324640
c 705	14	0.9	3282	4	US-09-315-793-51	Sequence 51, Appl
706	14	0.9	3312	1	US-08-049-473-1	Sequence 1, Appli
707	14	0.9	3312	1	US-08-312-648-1	Sequence 1, Appli
708	14	0.9	3312	5	PCT-US94-04190-1	Sequence 1, Appli
709	14	0.9	3389	1	US-08-395-580-1	Sequence 1, Appli
c 710	14	0.9	3396	4	US-08-974-549A-640	Sequence 640, App
c 711	14	0.9	3396	4	US-08-974-549A-641	Sequence 641, App
712	14	0.9	3426	1	US-08-205-018-1	Sequence 1, Appli
c 713	14	0.9	3505	1	US-08-454-097-34	Sequence 34, Appl
c 714	14	0.9	3505	1	US-08-468-036-43	Sequence 43, Appl
c 715	14	0.9	3505	2	US-08-376-843-43	Sequence 43, Appl
c 716	14	0.9	3505	3	US-08-185-359-34	Sequence 34, Appl
c 717	14	0.9	3568	4	US-09-103-510B-1	Sequence 1, Appli
718	14	0.9	3620	4	US-09-446-504-55	Sequence 55, Appl
c 719	14	0.9	3660	3	US-08-613-009A-5	Sequence 5, Appli
c 720	14	0.9	3672	1	US-08-491-357-1	Sequence 1, Appli
c 721	14	0.9	3672	3	US-08-968-633-1	Sequence 1, Appli
c 722	14	0.9	3672	3	US-09-196-466-1	Sequence 1, Appli
c 723	14	0.9	3672	5	PCT-US96-10823-1	Sequence 1, Appli
724	14	0.9	3727	1	US-08-249-380-1	Sequence 1, Appli
725	14	0.9	3827	2	US-08-447-031A-1	Sequence 1, Appli
726	14	0.9	3872	2	US-08-331-081B-4	Sequence 4, Appli
727	14	0.9	3879	1	US-08-530-010-1	Sequence 1, Appli
728	14	0.9	3879	1	US-08-484-101B-1	Sequence 1, Appli
c 729	14	0.9	3892	4	US-08-569-214-1	Sequence 1, Appli
c 730	14	0.9	3892	4	US-08-937-236-1	Sequence 1, Appli
c 731	14	0.9	3954	1	US-08-317-880-1	Sequence 1, Appli
c 732	14	0.9	3954	2	US-08-782-396-1	Sequence 1, Appli
c 733	14	0.9	3987	1	US-07-688-352C-19	Sequence 19, Appl
c 734	14	0.9	3987	2	US-08-474-379C-19	Sequence 19, Appl
c 735	14	0.9	3987	3	US-09-146-249A-19	Sequence 19, Appl
c 736	14	0.9	3987	3	US-08-206-188B-19	Sequence 19, Appl
c 737	14	0.9	3987	5	PCT-US91-02714-19	Sequence 19, Appl
738	14	0.9	4027	1	US-08-348-143-3	Sequence 3, Appli
739	14	0.9	4027	1	US-08-348-143-4	Sequence 4, Appli
740	14	0.9	4027	1	US-08-571-785-3	Sequence 3, Appli
741	14	0.9	4027	1	US-08-571-785-4	Sequence 4, Appli
c 742	14	0.9	4032	1	US-08-107-748-3	Sequence 3, Appli
c 743	14	0.9	4032	1	US-08-245-809-4	Sequence 4, Appli
c 744	14	0.9	4032	5	PCT-US92-01385-3	Sequence 3, Appli
745	14	0.9	4046	1	US-08-706-936-1	Sequence 1, Appli
c 746	14	0.9	4072	4	US-09-272-496-7	Sequence 7, Appli
c 747	14	0.9	4078	2	US-08-960-022-3	Sequence 3, Appli
c 748	14	0.9	4080	2	US-08-446-345-35	Sequence 35, Appl
c 749	14	0.9	4085	3	US-09-165-240-5	Sequence 5, Appli
750	14	0.9	4094	2	US-08-841-483-5	Sequence 5, Appli
751	14	0.9	4094	4	US-09-382-911-5	Sequence 5, Appli
c 752	14	0.9	4118	1	US-08-119-125A-3	Sequence 3, Appli
c 753	14	0.9	4157	1	US-07-807-043B-9	Sequence 9, Appli
c 754	14	0.9	4157	1	US-08-299-849B-9	Sequence 9, Appli
c 755	14	0.9	4157	2	US-08-142-368A-9	Sequence 9, Appli
c 756	14	0.9	4157	3	US-08-967-727-9	Sequence 9, Appli
c 757	14	0.9	4157	4	US-08-037-230D-9	Sequence 9, Appli
758	14	0.9	4164	1	US-08-188-582-3	Sequence 3, Appli
759	14	0.9	4164	1	US-08-646-715-3	Sequence 3, Appli
760	14	0.9	4255	5	PCT-US96-02331-14	Sequence 14, Appl
c 761	14	0.9	4291	2	US-08-417-210A-81	Sequence 81, Appl
762	14	0.9	4383	6	5175095-4	Patent No.5175095
763	14	0.9	4383	6	5177307-1	Patent No. 5177307
c 764	14	0.9	4512	1	US-08-224-391-52	Sequence 52, Appl
c 765	14	0.9	4512	1	US-08-484-304-52	Sequence 52, Appl
766	14	0.9	4612	2	US-08-447-031A-8	Sequence 8, Appli
767	14	0.9	4866	1	US-08-110-158-5	Sequence 5, Appli
c 768	14	0.9	4967	4	US-09-177-325-12	Sequence 12, Appl
769	14	0.9	5224	2	US-08-874-186-46	Sequence 46, Appl
c 770	14	0.9	5253	2	US-08-290-731C-3	Sequence 3, Appli
771	14	0.9	5275	1	US-08-485-588-1	Sequence 1, Appli
772	14	0.9	5275	1	US-08-484-565-1	Sequence 1, Appli
773	14	0.9	5275	2	US-08-480-751-1	Sequence 1, Appli
774	14	0.9	5275	2	US-08-943-986-1	Sequence 1, Appli
775	14	0.9	5275	3	US-08-353-784-1	Sequence 1, Appli
776	14	0.9	5275	3	US-08-484-719B-1	Sequence 1, Appli
c 777	14	0.9	5300	1	US-08-766-014-1	Sequence 1, Appli
778	14	0.9	5398	3	US-09-356-952-11	Sequence 11, Appl
c 779	14	0.9	5408	1	US-08-471-058-20	Sequence 20, Appl
c 780	14	0.9	5408	3	US-08-471-057-20	Sequence 20, Appl
c 781	14	0.9	5438	4	US-08-456-200B-5	Sequence 5, Appli
782	14	0.9	5475	2	US-08-680-327-1	Sequence 1, Appli
783	14	0.9	5475	4	US-09-228-246-3	Sequence 3, Appli
784	14	0.9	5503	2	US-08-726-012B-1	Sequence 1, Appli
785	14	0.9	5510	1	US-08-123-161A-7	Sequence 7, Appli
786	14	0.9	5510	1	US-08-483-278-7	Sequence 7, Appli
c 787	14	0.9	5855	1	US-08-592-214A-20	Sequence 20, Appl
c 788	14	0.9	5855	3	US-09-149-976-20	Sequence 20, Appl
c 789	14	0.9	5996	3	US-09-028-934-33	Sequence 33, Appl
790	14	0.9	6060	5	PCT-US96-09430-7	Sequence 7, Appli
c 791	14	0.9	6386	2	US-08-483-376-1	Sequence 1, Appli
c 792	14	0.9	6769	1	US-08-480-784-20	Sequence 20, Appl
c 793	14	0.9	6769	1	US-08-483-553-20	Sequence 20, Appl
c 794	14	0.9	6769	1	US-08-487-002-20	Sequence 20, Appl
c 795	14	0.9	6769	1	US-08-483-554B-20	Sequence 20, Appl
c 796	14	0.9	6769	1	US-08-488-011B-20	Sequence 20, Appl
c 797	14	0.9	6769	4	US-08-850-727-20	Sequence 20, Appl
c 798	14	0.9	6769	5	PCT-US95-10202-20	Sequence 20, Appl
c 799	14	0.9	6769	5	PCT-US95-10203-20	Sequence 20, Appl
c 800	14	0.9	6769	5	PCT-US95-10220-20	Sequence 20, Appl
c 801	14	0.9	6843	1	US-08-727-034-6	Sequence 6, Appli
802	14	0.9	6953	1	US-07-805-123C-2	Sequence 2, Appli
803	14	0.9	6953	1	US-08-033-081B-2	Sequence 2, Appli
c 804	14	0.9	6960	2	US-08-841-349-3	Sequence 3, Appli
c 805	14	0.9	6961	1	US-08-727-034-2	Sequence 2, Appli
c 806	14	0.9	7091	4	US-08-975-762-46	Sequence 46, Appl
c 807	14	0.9	7432	1	US-07-852-260-1	Sequence 1, Appli
c 808	14	0.9	7432	2	US-08-461-503-1	Sequence 1, Appli
c 809	14	0.9	7432	3	US-08-465-250-1	Sequence 1, Appli
810	14	0.9	7742	2	US-08-882-704A-4	Sequence 4, Appli
811	14	0.9	8060	4	US-08-766-528-1	Sequence 1, Appli
812	14	0.9	8132	4	US-08-766-528-3	Sequence 3, Appli
c 813	14	0.9	8176	2	US-08-841-349-5	Sequence 5, Appli
814	14	0.9	8779	2	US-08-750-703-4	Sequence 4, Appli
c 815	14	0.9	8878	1	US-08-206-176-3	Sequence 3, Appli
816	14	0.9	9060	4	US-08-378-313-20	Sequence 20, Appl
c 817	14	0.9	9103	1	US-08-466-033-182	Sequence 182, App
c 818	14	0.9	9103	1	US-08-444-733-182	Sequence 182, App
c 819	14	0.9	9103	2	US-08-464-134-182	Sequence 182, App
c 820	14	0.9	9103	2	US-08-461-361-182	Sequence 182, App
c 821	14	0.9	9103	2	US-08-485-910-182	Sequence 182, App
c 822	14	0.9	9103	5	PCT-US95-06266-156	Sequence 156, App
c 823	14	0.9	9511	1	US-07-925-695-6	Sequence 6, Appli
c 824	14	0.9	9511	1	US-07-925-695-7	Sequence 7, Appli
c 825	14	0.9	9636	1	US-08-323-170B-1	Sequence 1, Appli
c 826	14	0.9	10614	1	US-08-135-511-35	Sequence 35, Appl
c 827	14	0.9	10614	1	US-08-187-453-35	Sequence 35, Appl
828	14	0.9	10718	3	US-08-325-426B-1	Sequence 1, Appli
829	14	0.9	10952	1	US-08-602-036A-1	Sequence 1, Appli
830	14	0.9	10952	2	US-08-502-374A-1	Sequence 1, Appli

831	14	0.9	10952	2	US-08-642-407A-1	Sequence 1, Appli
832	14	0.9	10968	2	US-08-680-327-2	Sequence 2, Appli
833	14	0.9	10968	4	US-09-228-246-1	Sequence 1, Appli
834	14	0.9	11236	1	US-07-853-913-1	Sequence 1, Appli
c 835	14	0.9	11663	1	US-08-446-932-1	Sequence 1, Appli
c 836	14	0.9	11663	1	US-08-801-263A-1	Sequence 1, Appli
c 837	14	0.9	11663	1	US-08-801-263A-7	Sequence 7, Appli
c 838	14	0.9	11663	3	US-09-102-248-1	Sequence 1, Appli
c 839	14	0.9	11663	3	US-09-102-248-7	Sequence 7, Appli
c 840	14	0.9	11703	1	US-08-801-263A-8	Sequence 8, Appli
c 841	14	0.9	11703	3	US-09-102-248-8	Sequence 8, Appli
c 842	14	0.9	11717	1	US-08-801-263A-4	Sequence 4, Appli
c 843	14	0.9	11717	3	US-09-102-248-4	Sequence 4, Appli
844	14	0.9	11832	2	US-08-416-603-1	Sequence 1, Appli
845	14	0.9	12047	2	US-09-022-461-1	Sequence 1, Appli
c 846	14	0.9	12565	4	US-09-345-217-3	Sequence 3, Appli
c 847	14	0.9	12720	1	US-08-403-866-11	Sequence 11, Appl
c 848	14	0.9	13674	2	US-08-852-807-1	Sequence 1, Appli
c 849	14	0.9	13865	3	US-09-009-217-11	Sequence 11, Appl
c 850	14	0.9	13865	3	US-09-009-656-11	Sequence 11, Appl
c 851	14	0.9	13905	4	US-08-972-218-1	Sequence 1, Appli
c 852	14	0.9	14078	4	US-09-433-262-1	Sequence 1, Appli
c 853	14	0.9	14578	3	US-08-859-694-1	Sequence 1, Appli
c 854	14	0.9	15328	2	US-08-888-497-33	Sequence 33, Appl
c 855	14	0.9	15328	5	PCT-US94-07926-33	Sequence 33, Appl
c 856	14	0.9	16656	1	US-08-741-881-1	Sequence 1, Appli
c 857	14	0.9	16656	1	US-08-739-158-1	Sequence 1, Appli
c 858	14	0.9	16656	2	US-08-739-167-1	Sequence 1, Appli
c 859	14	0.9	16656	3	US-08-404-796-1	Sequence 1, Appli
c 860	14	0.9	16656	3	US-08-931-869-1	Sequence 1, Appli
c 861	14	0.9	17949	4	US-09-087-465-3	Sequence 3, Appli
862	14	0.9	18318	1	US-08-414-926A-6	Sequence 6, Appli
863	14	0.9	18318	2	US-08-926-922-6	Sequence 6, Appli
864	14	0.9	18318	3	US-09-253-682-6	Sequence 6, Appli
c 865	14	0.9	18596	4	US-09-318-448-11	Sequence 11, Appl
c 866	14	0.9	19011	1	US-08-310-356-36	Sequence 36, Appl
867	14	0.9	19307	3	US-08-836-022A-10	Sequence 10, Appl
868	14	0.9	19307	4	US-09-427-048A-10	Sequence 10, Appl
c 869	14	0.9	19557	5	PCT-US92-06300-1	Sequence 1, Appli
c 870	14	0.9	19932	2	US-08-477-451-25	Sequence 25, Appl
c 871	14	0.9	20303	1	US-08-370-975B-6	Sequence 6, Appli
c 872	14	0.9	24417	2	US-08-846-762-1	Sequence 1, Appli
c 873	14	0.9	24595	6	5428147-1	Patent No. 5428147
c 874	14	0.9	26764	1	US-08-370-975B-1	Sequence 1, Appli
c 875	14	0.9	36741	4	US-09-301-665-3	Sequence 3, Appli
876	14	0.9	38506	3	US-09-320-878-19	Sequence 19, Appl
877	14	0.9	40138	3	US-09-090-793-12	Sequence 12, Appl
c 878	14	0.9	43676	3	US-09-356-952-12	Sequence 12, Appl
c 879	14	0.9	49272	1	US-08-614-770A-1	Sequence 1, Appli
c 880	14	0.9	80161	3	US-09-036-987A-1	Sequence 1, Appli
881	14	0.9	176373	3	US-09-128-155-17	Sequence 17, Appl
882	14	0.9	246240	2	US-08-724-394A-20	Sequence 20, Appl
c 883	14	0.9	246240	2	US-08-724-394A-20	Sequence 20, Appl
884	14	0.9	246240	2	US-08-724-394A-21	Sequence 21, Appl
c 885	14	0.9	246240	2	US-08-724-394A-21	Sequence 21, Appl
886	14	0.9	246240	2	US-08-724-394A-22	Sequence 22, Appl
c 887	14	0.9	246240	2	US-08-724-394A-22	Sequence 22, Appl
c 888	13	0.8	15	1	US-08-319-492B-18	Sequence 18, Appl
889	13	0.8	16	3	US-09-079-839-1	Sequence 1, Appli
c 890	13	0.8	17	1	US-08-390-850-423	Sequence 423, App
c 891	13	0.8	17	1	US-08-390-850-424	Sequence 424, App
c 892	13	0.8	17	1	US-08-435-634-423	Sequence 423, App
c 893	13	0.8	17	1	US-08-435-634-424	Sequence 424, App
c 894	13	0.8	18	2	US-08-823-516-161	Sequence 161, App
895	13	0.8	19	3	US-08-846-020A-21	Sequence 21, Appl
c 896	13	0.8	20	2	US-07-963-538B-13	Sequence 13, Appl
897	13	0.8	20	2	US-07-963-538B-29	Sequence 29, Appl
c 898	13	0.8	20	2	US-08-730-876-2	Sequence 2, Appli
899	13	0.8	20	2	US-09-289-368-63	Sequence 63, Appl
c 900	13	0.8	21	1	US-08-468-580-40	Sequence 40, Appl
901	13	0.8	21	1	US-08-468-580-45	Sequence 45, Appl
c 902	13	0.8	21	4	US-08-643-212-62	Sequence 62, Appl
903	13	0.8	21	4	US-08-643-212-67	Sequence 67, Appl
c 904	13	0.8	21	5	PCT-US95-03731-40	Sequence 40, Appl
905	13	0.8	21	5	PCT-US95-03731-45	Sequence 45, Appl
c 906	13	0.8	23	1	US-08-316-293-14	Sequence 14, Appl
907	13	0.8	23	3	US-08-691-045-37	Sequence 37, Appl
c 908	13	0.8	24	1	US-07-741-940-55	Sequence 55, Appl
c 909	13	0.8	24	1	US-08-289-548A-55	Sequence 55, Appl
c 910	13	0.8	24	1	US-08-452-654-55	Sequence 55, Appl
c 911	13	0.8	24	1	US-08-452-655B-55	Sequence 55, Appl
c 912	13	0.8	24	1	US-08-465-590-52	Sequence 52, Appl
c 913	13	0.8	24	2	US-08-508-088-5	Sequence 5, Appli
c 914	13	0.8	24	2	US-08-674-149A-11	Sequence 11, Appl
c 915	13	0.8	24	2	US-09-009-925-5	Sequence 5, Appli
916	13	0.8	24	3	US-08-707-743-4	Sequence 4, Appli
c 917	13	0.8	24	3	US-08-450-582-55	Sequence 55, Appl
918	13	0.8	24	4	US-08-283-300A-20	Sequence 20, Appl
c 919	13	0.8	24	4	US-09-311-260-189	Sequence 189, App
c 920	13	0.8	24	4	US-08-711-417C-52	Sequence 52, Appl
c 921	13	0.8	24	5	PCT-US93-08743-52	Sequence 52, Appl
922	13	0.8	24	5	PCT-US95-09345-20	Sequence 20, Appl
c 923	13	0.8	25	1	US-08-125-012-19	Sequence 19, Appl
924	13	0.8	25	1	US-08-125-012-20	Sequence 20, Appl
c 925	13	0.8	25	1	US-08-221-579A-13	Sequence 13, Appl
926	13	0.8	25	1	US-08-221-579A-14	Sequence 14, Appl
c 927	13	0.8	25	1	US-08-783-818-19	Sequence 19, Appl
928	13	0.8	25	1	US-08-783-818-20	Sequence 20, Appl
c 929	13	0.8	25	2	US-08-453-349-19	Sequence 19, Appl
930	13	0.8	25	2	US-08-453-349-20	Sequence 20, Appl
931	13	0.8	25	2	US-08-273-402B-1	Sequence 1, Appli
c 932	13	0.8	25	2	US-08-704-701-13	Sequence 13, Appl
933	13	0.8	25	2	US-08-704-701-14	Sequence 14, Appl
c 934	13	0.8	25	2	US-08-979-385B-17	Sequence 17, Appl
935	13	0.8	25	2	US-08-979-385B-18	Sequence 18, Appl
c 936	13	0.8	25	4	US-09-321-461-13	Sequence 13, Appl
937	13	0.8	25	4	US-09-321-461-14	Sequence 14, Appl
c 938	13	0.8	25	5	PCT-US95-16206A-5	Sequence 5, Appli
939	13	0.8	25	5	PCT-US95-16206A-6	Sequence 6, Appli
940	13	0.8	26	3	US-08-582-740-4	Sequence 4, Appli
c 941	13	0.8	26	4	US-08-961-810-55	Sequence 55, Appl
c 942	13	0.8	26	4	US-08-961-810-94	Sequence 94, Appl
c 943	13	0.8	26	4	US-08-352-902D-55	Sequence 55, Appl
c 944	13	0.8	26	4	US-08-352-902D-94	Sequence 94, Appl
c 945	13	0.8	27	3	US-08-998-099-166	Sequence 166, App
946	13	0.8	28	2	US-08-317-310A-14	Sequence 14, Appl
947	13	0.8	28	5	PCT-US95-13041-14	Sequence 14, Appl
c 948	13	0.8	29	1	US-08-647-584-50	Sequence 50, Appl
949	13	0.8	30	1	US-08-201-697-14	Sequence 14, Appl
c 950	13	0.8	30	1	US-08-468-557-25	Sequence 25, Appl
c 951	13	0.8	30	1	US-08-451-715A-64	Sequence 64, Appl
c 952	13	0.8	30	1	US-08-493-638-2	Sequence 2, Appli
c 953	13	0.8	30	2	US-08-743-130A-36	Sequence 36, Appl
c 954	13	0.8	30	2	US-08-428-257A-33	Sequence 33, Appl
c 955	13	0.8	30	2	US-08-742-026-16	Sequence 16, Appl
c 956	13	0.8	30	2	US-08-415-593-24	Sequence 24, Appl
c 957	13	0.8	30	2	US-08-962-284-21	Sequence 21, Appl
c 958	13	0.8	32	4	US-08-729-601A-81	Sequence 81, Appl
c 959	13	0.8	33	1	US-08-250-310-9	Sequence 9, Appli
c 960	13	0.8	33	1	US-08-439-404-9	Sequence 9, Appli
961	13	0.8	33	2	US-08-709-874A-23	Sequence 23, Appl
c 962	13	0.8	33	3	US-08-839-327-9	Sequence 9, Appli
c 963	13	0.8	33	3	US-09-340-781B-9	Sequence 9, Appli
964	13	0.8	33	4	US-09-104-382-23	Sequence 23, Appl
c 965	13	0.8	33	6	5519127-18	Patent No. 5519127
c 966	13	0.8	34	2	US-08-962-284-14	Sequence 14, Appl
967	13	0.8	34	3	US-08-813-574-5	Sequence 5, Appli
c 968	13	0.8	36	1	US-08-334-847-801	Sequence 801, App
c 969	13	0.8	36	1	US-08-311-486C-335	Sequence 335, App
c 970	13	0.8	38	1	US-08-373-124A-1167	Sequence 1167, Ap
c 971	13	0.8	38	1	US-08-435-628-1167	Sequence 1167, Ap
c 972	13	0.8	39	6	5519127-32	Patent No. 5519127
973	13	0.8	40	1	US-08-105-483-161	Sequence 161, App
974	13	0.8	40	1	US-08-709-209-161	Sequence 161, App
975	13	0.8	40	1	US-08-303-275-49	Sequence 49, Appl
976	13	0.8	40	1	US-08-458-101-161	Sequence 161, App

977	13	0.8	40	2	US-08-687-916-9	Sequence 9, Appli
978	13	0.8	40	4	US-09-138-614-9	Sequence 9, Appli
c 979	13	0.8	41	1	US-07-931-473B-165	Sequence 165, App
c 980	13	0.8	41	1	US-07-714-131C-165	Sequence 165, App
c 981	13	0.8	41	1	US-08-412-110-165	Sequence 165, App
c 982	13	0.8	41	1	US-08-409-442A-165	Sequence 165, App
c 983	13	0.8	41	1	US-08-399-696-43	Sequence 43, Appl
c 984	13	0.8	41	2	US-08-469-609A-165	Sequence 165, App
c 985	13	0.8	41	3	US-09-143-190-165	Sequence 165, App
986	13	0.8	43	2	US-08-988-128-7	Sequence 7, Appli
987	13	0.8	43	2	US-08-988-128-8	Sequence 8, Appli
c 988	13	0.8	44	1	US-07-931-473B-242	Sequence 242, App
c 989	13	0.8	44	1	US-07-995-950-4	Sequence 4, Appli
c 990	13	0.8	44	1	US-07-714-131C-242	Sequence 242, App
c 991	13	0.8	44	1	US-08-344-695-4	Sequence 4, Appli
c 992	13	0.8	44	1	US-08-412-110-242	Sequence 242, App
c 993	13	0.8	44	1	US-08-409-442A-242	Sequence 242, App
c 994	13	0.8	44	1	US-08-300-582-4	Sequence 4, Appli
c 995	13	0.8	44	2	US-08-469-609A-242	Sequence 242, App
c 996	13	0.8	44	3	US-09-143-190-242	Sequence 242, App
997	13	0.8	47	1	US-08-171-389-314	Sequence 314, App
998	13	0.8	47	1	US-08-123-936-314	Sequence 314, App
999	13	0.8	47	2	US-08-475-228A-314	Sequence 314, App
1000	13	0.8	47	3	US-08-482-080A-314	Sequence 314, App

ALIGNMENTS

```
RESULT   1
US-08-637-759B-24
; Sequence 24, Application US/08637759B
; Patent No. 5876931
;  GENERAL INFORMATION:
;    APPLICANT:  David William Holden
;    TITLE OF INVENTION:  Identification of Genes
;    NUMBER OF SEQUENCES:  501
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/637,759B
;      FILING DATE:  03-MAY-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02875
;      FILING DATE:  11-DEC-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  RPMS 101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  303 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Partial sequence od Salmonella typhimurium
;      ORGANISM:  virulence gene
US-08-637-759B-24


  Query Match             1.1%;  Score 17;  DB 2;  Length 303;
  Best Local Similarity   100.0%;  Pred. No. 39;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      63 tgcagatattcccgggt 79
           |||||||||||||||||
Db     260 TGCAGATATTCCCGGGT 276


RESULT   2
US-08-871-355A-24
; Sequence 24, Application US/08871355A
; Patent No. 6015669
;  GENERAL INFORMATION:
;    APPLICANT:  David William Holden
;    TITLE OF INVENTION:  Identification of Genes
;    NUMBER OF SEQUENCES:  501
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/871,355A
;      FILING DATE:  09-JUN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02875
;      FILING DATE:  11-DEC-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  RPMS 101 CON
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  303 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Partial sequence of Salmonella typhimurium
;      ORGANISM:  virulence gene
US-08-871-355A-24


  Query Match             1.1%;  Score 17;  DB 3;  Length 303;
  Best Local Similarity   100.0%;  Pred. No. 39;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      63 tgcagatattcccgggt 79
           |||||||||||||||||
```

```
Db    260 TGCAGATATTCCCGGGT 276


RESULT   3
US-08-637-759B-38/c
; Sequence 38, Application US/08637759B
; Patent No. 5876931
;  GENERAL INFORMATION:
;    APPLICANT:  David William Holden
;    TITLE OF INVENTION:  Identification of Genes
;    NUMBER OF SEQUENCES:  501
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/637,759B
;      FILING DATE:  03-MAY-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02875
;      FILING DATE:  11-DEC-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  RPMS 101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5746 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  DNA sequence of VGC II cluster C
US-08-637-759B-38


  Query Match             1.1%;  Score 17;  DB 2;  Length 5746;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      63 tgcagatattcccgggt 79
           |||||||||||||||||
Db     212 TGCAGATATTCCCGGGT 196


RESULT   4
US-08-871-355A-38/c
; Sequence 38, Application US/08871355A
; Patent No. 6015669
;  GENERAL INFORMATION:
;    APPLICANT:  David William Holden
;    TITLE OF INVENTION:  Identification of Genes
;    NUMBER OF SEQUENCES:  501
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/871,355A
;      FILING DATE:  09-JUN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02875
;      FILING DATE:  11-DEC-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  RPMS 101 CON
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5746 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  DNA sequence of VGC II cluster C
US-08-871-355A-38


  Query Match             1.1%;  Score 17;  DB 3;  Length 5746;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      63 tgcagatattcccgggt 79
           |||||||||||||||||
Db     212 TGCAGATATTCCCGGGT 196


RESULT   5
US-08-243-542-9/c
; Sequence 9, Application US/08243542
; Patent No. 5552526
;  GENERAL INFORMATION:
;    APPLICANT:  NAKAMURA, YUSUKE
;    APPLICANT:  : EMI, MITSURU
;    TITLE OF INVENTION:  MDC PROTEINS AND DNAs
;    TITLE OF INVENTION:  ENCODING THE SAME
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLYNN, THIEL, BOUTELL & TANIS P.C.
;      STREET:  2026 Rambling Road
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49008-1699
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inches, 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.0
;      SOFTWARE:  WordPerfect 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/243,542
;      FILING DATE:
```

```
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP 5-136602
;       FILING DATE:  14 MAY 1993
;       APPLICATION NUMBER:  JP 5-257455
;       FILING DATE:  22 SEPTEMBER 1993
;       APPLICATION NUMBER:  JP 6-49904
;       FILING DATE:  23 FEBRUARY 1994
;       APPLICATION NUMBER:  JP 6-73328
;       FILING DATE:  12 APRIL 1994
;       APPLICATION NUMBER:  JP 6-84470
;       FILING DATE:  22 APRIL 1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Terryence F. Chapman
;       REGISTRATION NUMBER:  32 549
;       REFERENCE/DOCKET NUMBER:  Furuya Case 1313
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (616) 381-1156
;       TELEFAX:  (616) 381-5465
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9278 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Genomic DNA
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
;     IMMEDIATE SOURCE:
;       LIBRARY:  human DNA cosmid library
;     FEATURE:
;       NAME/KEY:  exon 1
;       LOCATION:  28..44
;     FEATURE:
;       NAME/KEY:  exon 2
;       LOCATION:  308..374
;     FEATURE:
;       NAME/KEY:  exon 3
;       LOCATION:  909..994
;     FEATURE:
;       NAME/KEY:  exon 4
;       LOCATION:  1081..1156
;     FEATURE:
;       NAME/KEY:  exon 5
;       LOCATION:  1591..1657
;     FEATURE:
;       NAME/KEY:  exon 6
;       LOCATION:  1725..1792
;     FEATURE:
;       NAME/KEY:  exon 7
;       LOCATION:  2182..2256
;     FEATURE:
;       NAME/KEY:  exon 8
;       LOCATION:  2339..2410
;     FEATURE:
;       NAME/KEY:  exon 9
;       LOCATION:  2588..2754
;     FEATURE:
;       NAME/KEY:  exon 10
;       LOCATION:  3248..3332
;     FEATURE:
;       NAME/KEY:  exon 11
;       LOCATION:  3445..3535
;     FEATURE:
;       NAME/KEY:  exon 12
;       LOCATION:  3645..3696
;     FEATURE:
;       NAME/KEY:  exon 13
;       LOCATION:  4014..4113
;     FEATURE:
;       NAME/KEY:  exon 14
;       LOCATION:  4196..4267
;     FEATURE:
;       NAME/KEY:  exon 15
;       LOCATION:  4386..4478
;     FEATURE:
;       NAME/KEY:  exon 16
;       LOCATION:  4920..5000
;     FEATURE:
;       NAME/KEY:  exon 17
;       LOCATION:  5347..5397
;     FEATURE:
;       NAME/KEY:  exon 18
;       LOCATION:  5501..5564
;     FEATURE:
;       NAME/KEY:  exon 19
;       LOCATION:  5767..5866
;     FEATURE:
;       NAME/KEY:  exon 20
;       LOCATION:  6073..6202
;     FEATURE:
;       NAME/KEY:  exon 21
;       LOCATION:  6300..6468
;     FEATURE:
;       NAME/KEY:  exon 22
;       LOCATION:  6557..6671
;     FEATURE:
;       NAME/KEY:  exon 23
;       LOCATION:  6756..6846
;     FEATURE:
;       NAME/KEY:  exon 24
;       LOCATION:  7829..7846
;     FEATURE:
;       NAME/KEY:  exon 25
;       LOCATION:  8165..9038
US-08-243-542-9


  Query Match             1.1%;  Score 17;  DB 1;  Length 9278;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1232 atgaggaggagggaagg 1248
          |||||||||||||||||
Db   2448 ATGAGGAGGAGGGAAGG 2432


RESULT   6
US-08-477-407-9/c
; Sequence 9, Application US/08477407
; Patent No. 5631351
;  GENERAL INFORMATION:
;    APPLICANT:  NAKAMURA, YUSUKE
;    APPLICANT:  :  EMI, MITSURU
;    TITLE OF INVENTION:  MDC PROTEINS AND DNAs
;    TITLE OF INVENTION:  ENCODING THE SAME
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLYNN, THIEL, BOUTELL & TANIS P.C.
;      STREET:  2026 Rambling Road
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49008-1699
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inches, 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.0
;      SOFTWARE:  WordPerfect 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,407
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/243,542
;      FILING DATE:  13-MAY-1994
```

```
;      APPLICATION NUMBER:  JP 5-136602
;      FILING DATE:  14 MAY 1993
;      APPLICATION NUMBER:  JP 5-257455
;      FILING DATE:  22 SEPTEMBER 1993
;      APPLICATION NUMBER:  JP 6-49904
;      FILING DATE:  23 FEBRUARY 1994
;      APPLICATION NUMBER:  JP 6-73328
;      FILING DATE:  12 APRIL 1994
;      APPLICATION NUMBER:  JP 6-84470
;      FILING DATE:  22 APRIL 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terryence F. Chapman
;      REGISTRATION NUMBER:  32 549
;      REFERENCE/DOCKET NUMBER:  Furuya Case 1313
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (616) 381-1156
;      TELEFAX:  (616) 381-5465
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9278 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    IMMEDIATE SOURCE:
;      LIBRARY:  human DNA cosmid library
;    FEATURE:
;      NAME/KEY:  exon 1
;      LOCATION:  28..44
;    FEATURE:
;      NAME/KEY:  exon 2
;      LOCATION:  308..374
;    FEATURE:
;      NAME/KEY:  exon 3
;      LOCATION:  909..994
;    FEATURE:
;      NAME/KEY:  exon 4
;      LOCATION:  1081..1156
;    FEATURE:
;      NAME/KEY:  exon 5
;      LOCATION:  1591..1657
;    FEATURE:
;      NAME/KEY:  exon 6
;      LOCATION:  1725..1792
;    FEATURE:
;      NAME/KEY:  exon 7
;      LOCATION:  2182..2256
;    FEATURE:
;      NAME/KEY:  exon 8
;      LOCATION:  2339..2410
;    FEATURE:
;      NAME/KEY:  exon 9
;      LOCATION:  2588..2754
;    FEATURE:
;      NAME/KEY:  exon 10
;      LOCATION:  3248..3332
;    FEATURE:
;      NAME/KEY:  exon 11
;      LOCATION:  3445..3535
;    FEATURE:
;      NAME/KEY:  exon 12
;      LOCATION:  3645..3696
;    FEATURE:
;      NAME/KEY:  exon 13
;      LOCATION:  4014..4113
;    FEATURE:
;      NAME/KEY:  exon 14
;      LOCATION:  4196..4267
;    FEATURE:
;      NAME/KEY:  exon 15
;      LOCATION:  4386..4478
;    FEATURE:
;      NAME/KEY:  exon 16
;      LOCATION:  4920..5000
;    FEATURE:
;      NAME/KEY:  exon 17
;      LOCATION:  5347..5397
;    FEATURE:
;      NAME/KEY:  exon 18
;      LOCATION:  5501..5564
;    FEATURE:
;      NAME/KEY:  exon 19
;      LOCATION:  5767..5866
;    FEATURE:
;      NAME/KEY:  exon 20
;      LOCATION:  6073..6202
;    FEATURE:
;      NAME/KEY:  exon 21
;      LOCATION:  6300..6468
;    FEATURE:
;      NAME/KEY:  exon 22
;      LOCATION:  6557..6671
;    FEATURE:
;      NAME/KEY:  exon 23
;      LOCATION:  6756..6846
;    FEATURE:
;      NAME/KEY:  exon 24
;      LOCATION:  7829..7846
;    FEATURE:
;      NAME/KEY:  exon 25
;      LOCATION:  8165..9038
US-08-477-407-9

  Query Match             1.1%;  Score 17;  DB 1;  Length 9278;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1232 atgaggaggagggaagg 1248
           ||||||||||||||||||
Db    2448 ATGAGGAGGAGGGAAGG 2432


RESULT   7
US-08-484-355-9/c
; Sequence 9, Application US/08484355
; Patent No. 5705341
;  GENERAL INFORMATION:
;    APPLICANT:  NAKAMURA, YUSUKE
;    APPLICANT:  :  EMI, MITSURU
;    TITLE OF INVENTION:  MDC PROTEINS AND DNAs
;    TITLE OF INVENTION:  ENCODING THE SAME
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLYNN, THIEL, BOUTELL & TANIS P.C.
;      STREET:  2026 Rambling Road
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49008-1699
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inches, 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.0
;      SOFTWARE:  WordPerfect 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,355
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/243,542
;      FILING DATE:  13-MAY-1994
;      APPLICATION NUMBER:  JP 5-136602
;      FILING DATE:  14 MAY 1993
```

```
;      APPLICATION NUMBER:  JP 5-257455
;      FILING DATE:  22 SEPTEMBER 1993
;      APPLICATION NUMBER:  JP 6-49904
;      FILING DATE:  23 FEBRUARY 1994
;      APPLICATION NUMBER:  JP 6-73328
;      FILING DATE:  12 APRIL 1994
;      APPLICATION NUMBER:  JP 6-84470
;      FILING DATE:  22 APRIL 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terryence F. Chapman
;      REGISTRATION NUMBER:  32 549
;      REFERENCE/DOCKET NUMBER:  Furuya Case 1313
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (616) 381-1156
;      TELEFAX:  (616) 381-5465
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9278 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    IMMEDIATE SOURCE:
;      LIBRARY:  human DNA cosmid library
;    FEATURE:
;      NAME/KEY:  exon 1
;      LOCATION:  28..44
;    FEATURE:
;      NAME/KEY:  exon 2
;      LOCATION:  308..374
;    FEATURE:
;      NAME/KEY:  exon 3
;      LOCATION:  909..994
;    FEATURE:
;      NAME/KEY:  exon 4
;      LOCATION:  1081..1156
;    FEATURE:
;      NAME/KEY:  exon 5
;      LOCATION:  1591..1657
;    FEATURE:
;      NAME/KEY:  exon 6
;      LOCATION:  1725..1792
;    FEATURE:
;      NAME/KEY:  exon 7
;      LOCATION:  2182..2256
;    FEATURE:
;      NAME/KEY:  exon 8
;      LOCATION:  2339..2410
;    FEATURE:
;      NAME/KEY:  exon 9
;      LOCATION:  2588..2754
;    FEATURE:
;      NAME/KEY:  exon 10
;      LOCATION:  3248..3332
;    FEATURE:
;      NAME/KEY:  exon 11
;      LOCATION:  3445..3535
;    FEATURE:
;      NAME/KEY:  exon 12
;      LOCATION:  3645..3696
;    FEATURE:
;      NAME/KEY:  exon 13
;      LOCATION:  4014..4113
;    FEATURE:
;      NAME/KEY:  exon 14
;      LOCATION:  4196..4267
;    FEATURE:
;      NAME/KEY:  exon 15
;      LOCATION:  4386..4478
;    FEATURE:
;      NAME/KEY:  exon 16
;      LOCATION:  4920..5000
;    FEATURE:
;      NAME/KEY:  exon 17
;      LOCATION:  5347..5397
;    FEATURE:
;      NAME/KEY:  exon 18
;      LOCATION:  5501..5564
;    FEATURE:
;      NAME/KEY:  exon 19
;      LOCATION:  5767..5866
;    FEATURE:
;      NAME/KEY:  exon 20
;      LOCATION:  6073..6202
;    FEATURE:
;      NAME/KEY:  exon 21
;      LOCATION:  6300..6468
;    FEATURE:
;      NAME/KEY:  exon 22
;      LOCATION:  6557..6671
;    FEATURE:
;      NAME/KEY:  exon 23
;      LOCATION:  6756..6846
;    FEATURE:
;      NAME/KEY:  exon 24
;      LOCATION:  7829..7846
;    FEATURE:
;      NAME/KEY:  exon 25
;      LOCATION:  8165..9038
US-08-484-355-9


  Query Match             1.1%;  Score 17;  DB 1;  Length 9278;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1232 atgaggaggagggaagg 1248
          |||||||||||||||||
Db   2448 ATGAGGAGGAGGGAAGG 2432


RESULT  8
US-08-345-321-5/c
; Sequence 5, Application US/08345321
; Patent No. 5914109
;  GENERAL INFORMATION:
;    APPLICANT:  ZOLLA-PAZNER, Susan
;    APPLICANT:  GORNY, Miroslav K.
;    TITLE OF INVENTION:  HETEROHYBRIDOMAS PRODUCING HUMAN
;    TITLE OF INVENTION:  MONOCLONAL ANTIBODIES TO HIV-1
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/345,321
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/872,675
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Browdy, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  ZOLLA-PAZNER1B
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  450 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..450
US-08-345-321-5


  Query Match             1.0%;  Score 16;  DB 2;  Length 450;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      29 cagttctgcttttaat 44
           ||||||||||||||||
Db     223 CAGTTCTGCTTTTAAT 208


RESULT   9
US-08-491-988-25
; Sequence 25, Application US/08491988
; Patent No. 5973116
;  GENERAL INFORMATION:
;    APPLICANT:  EPENETOS, AGAMEMNON A.
;    APPLICANT:  SPOONER, ROBERT A.
;    APPLICANT:  DEONARAIN, MAHENDRA
;    TITLE OF INVENTION:  Compounds for targeting
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAULAY NISSEN GOLDBERG KIEL & HAND, LLP
;      STREET:  261 MADISON AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,988
;      FILING DATE:  18-DEC-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDBERG, JULES E.
;      REGISTRATION NUMBER:  24,408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-986-4090
;      TELEFAX:  212-818-9479
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1178 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-491-988-25


  Query Match             1.0%;  Score 16;  DB 2;  Length 1178;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1508 gcagcacatggactct 1523
           ||||||||||||||||
Db     822 GCAGCACATGGACTCT 837


RESULT  10
US-08-491-988-28
; Sequence 28, Application US/08491988
; Patent No. 5973116
;  GENERAL INFORMATION:
;    APPLICANT:  EPENETOS, AGAMEMNON A.
;    APPLICANT:  SPOONER, ROBERT A.
;    APPLICANT:  DEONARAIN, MAHENDRA
;    TITLE OF INVENTION:  Compounds for targeting
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAULAY NISSEN GOLDBERG KIEL & HAND, LLP
;      STREET:  261 MADISON AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,988
;      FILING DATE:  18-DEC-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDBERG, JULES E.
;      REGISTRATION NUMBER:  24,408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-986-4090
;      TELEFAX:  212-818-9479
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1178 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-491-988-28


  Query Match             1.0%;  Score 16;  DB 2;  Length 1178;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1508 gcagcacatggactct 1523
           ||||||||||||||||
Db     822 GCAGCACATGGACTCT 837


RESULT  11
US-08-491-988-27
; Sequence 27, Application US/08491988
; Patent No. 5973116
;  GENERAL INFORMATION:
;    APPLICANT:  EPENETOS, AGAMEMNON A.
;    APPLICANT:  SPOONER, ROBERT A.
;    APPLICANT:  DEONARAIN, MAHENDRA
;    TITLE OF INVENTION:  Compounds for targeting
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAULAY NISSEN GOLDBERG KIEL & HAND, LLP
;      STREET:  261 MADISON AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
```

```
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,988
;      FILING DATE:  18-DEC-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDBERG, JULES E.
;      REGISTRATION NUMBER:  24,408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-986-4090
;      TELEFAX:  212-818-9479
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1202 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-491-988-27


  Query Match             1.0%;  Score 16;  DB 2;  Length 1202;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1508 gcagcacatggactct 1523
          ||||||||||||||||
Db    846 GCAGCACATGGACTCT 861


RESULT  12
US-08-491-988-12
; Sequence 12, Application US/08491988
; Patent No. 5973116
;  GENERAL INFORMATION:
;    APPLICANT:  EPENETOS, AGAMEMNON A.
;    APPLICANT:  SPOONER, ROBERT A.
;    APPLICANT:  DEONARAIN, MAHENDRA
;    TITLE OF INVENTION:  Compounds for targeting
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAULAY NISSEN GOLDBERG KIEL & HAND, LLP
;      STREET:  261 MADISON AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,988
;      FILING DATE:  18-DEC-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDBERG, JULES E.
;      REGISTRATION NUMBER:  24,408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-986-4090
;      TELEFAX:  212-818-9479
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1226 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-491-988-12


  Query Match             1.0%;  Score 16;  DB 2;  Length 1226;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1508 gcagcacatggactct 1523
          ||||||||||||||||
Db    852 GCAGCACATGGACTCT 867


RESULT  13
US-08-491-988-11
; Sequence 11, Application US/08491988
; Patent No. 5973116
;  GENERAL INFORMATION:
;    APPLICANT:  EPENETOS, AGAMEMNON A.
;    APPLICANT:  SPOONER, ROBERT A.
;    APPLICANT:  DEONARAIN, MAHENDRA
;    TITLE OF INVENTION:  Compounds for targeting
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAULAY NISSEN GOLDBERG KIEL & HAND, LLP
;      STREET:  261 MADISON AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,988
;      FILING DATE:  18-DEC-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDBERG, JULES E.
;      REGISTRATION NUMBER:  24,408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-986-4090
;      TELEFAX:  212-818-9479
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1235 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-491-988-11


  Query Match             1.0%;  Score 16;  DB 2;  Length 1235;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1508 gcagcacatggactct 1523
          ||||||||||||||||
Db    861 GCAGCACATGGACTCT 876


RESULT  14
US-08-491-988-10
; Sequence 10, Application US/08491988
; Patent No. 5973116
;  GENERAL INFORMATION:
;    APPLICANT:  EPENETOS, AGAMEMNON A.
```

```
;    APPLICANT:  SPOONER, ROBERT A.
;    APPLICANT:  DEONARAIN, MAHENDRA
;    TITLE OF INVENTION:  Compounds for targeting
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAULAY NISSEN GOLDBERG KIEL & HAND, LLP
;      STREET:  261 MADISON AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,988
;      FILING DATE:  18-DEC-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDBERG, JULES E.
;      REGISTRATION NUMBER:  24,408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-986-4090
;      TELEFAX:  212-818-9479
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1259 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-491-988-10


  Query Match             1.0%;  Score 16;  DB 2;  Length 1259;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1508 gcagcacatggactct 1523
          ||||||||||||||||
Db    885 GCAGCACATGGACTCT 900


RESULT  15
US-08-491-988-24
; Sequence 24, Application US/08491988
; Patent No. 5973116
;  GENERAL INFORMATION:
;    APPLICANT:  EPENETOS, AGAMEMNON A.
;    APPLICANT:  SPOONER, ROBERT A.
;    APPLICANT:  DEONARAIN, MAHENDRA
;    TITLE OF INVENTION:  Compounds for targeting
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAULAY NISSEN GOLDBERG KIEL & HAND, LLP
;      STREET:  261 MADISON AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,988
;      FILING DATE:  18-DEC-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDBERG, JULES E.
;      REGISTRATION NUMBER:  24,408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-986-4090
;      TELEFAX:  212-818-9479
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1259 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-08-491-988-24


  Query Match             1.0%;  Score 16;  DB 2;  Length 1259;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1508 gcagcacatggactct 1523
          ||||||||||||||||
Db    915 GCAGCACATGGACTCT 930


Search completed: October 30, 2001, 12:29:22
Job time: 2734 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2001, 10:53:43 ; Search time 1191.07 Seconds
(without alignments)
12190.369 Million cell updates/sec

Title: US-09-619-049-854
Perfect score: 1536
Sequence: 1 atgaggcttccttgggtgat.....ggactctaatgatggattct 1536

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 10228115 seqs, 4726426750 residues

Word size : 0

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est75:*
137: gb_est76:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est67:*
161: gb_est68:*
162: gb_est69:*
163: gb_est70:*
164: gb_est71:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est91:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthum29:*
177: em_esthum30:*
178: em_esthum31:*
179: em_esthum32:*
180: em_esthum33:*
181: em_estom3:*
182: em_estpl11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*
190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	31	2.0	559	223	AQ026293	AQ026293 1(3)j1C7
c	2	22	1.4	904	140	BE781852	BE781852 601470173
c	3	21	1.4	104	112	AW176349	AW176349 QV0-BT023
	4	21	1.4	229	24	AI752336	AI752336 cn15a10.x
	5	21	1.4	313	7	AA457068	AA457068 aa90f05.r
	6	21	1.4	590	166	BE357901	BE357901 DG1_23_H0
	7	21	1.4	820	150	BF574583	BF574583 602133237
c	8	20	1.3	250	161	BB582897	BB582897 BB582897
c	9	20	1.3	353	2	AA138114	AA138114 mq75d06.r
c	10	20	1.3	380	106	AL588417	AL588417 AL588417
	11	20	1.3	497	231	AQ667727	AQ667727 HS_2137_B
c	12	20	1.3	518	147	BF318468	BF318468 ux10c05.x
c	13	20	1.3	536	248	AZ739776	AZ739776 RPCI-24-1
c	14	20	1.3	585	173	BG081170	BG081170 H3061H07-
c	15	20	1.3	591	248	AZ739770	AZ739770 RPCI-24-1
c	16	20	1.3	627	105	AL121387	AL121387 DKFZp762E
	17	20	1.3	643	173	BG068113	BG068113 H3061H07-
c	18	20	1.3	919	192	AK017892	AK017892 Mus muscu
c	19	20	1.3	1534	142	BE972563	BE972563 601651944
c	20	20	1.3	1563	145	BF183692	BF183692 601809689
	21	19	1.2	101	159	N67195	N67195 yz48a02.s1
c	22	19	1.2	110	112	AW176740	AW176740 RC0-CT008
	23	19	1.2	117	22	AI583414	AI583414 ts09d11.x
	24	19	1.2	120	116	AW496804	AW496804 ERaD3g5_1
	25	19	1.2	122	21	AI538646	AI538646 tp70f09.x
	26	19	1.2	124	11	AA709093	AA709093 zf96h09.s
	27	19	1.2	129	21	AI520743	AI520743 ti07a05.x
	28	19	1.2	130	114	AW272321	AW272321 xu17h05.x
	29	19	1.2	137	102	AI858414	AI858414 wl38a06.x
	30	19	1.2	138	12	AA853127	AA853127 NHTBCae03
	31	19	1.2	144	12	AA852124	AA852124 NHTBCae10
c	32	19	1.2	145	121	AW836339	AW836339 CM4-LT002
c	33	19	1.2	146	110	AW022210	AW022210 df34d07.y
	34	19	1.2	148	159	N66903	N66903 yz57h02.s1
	35	19	1.2	148	173	BG056595	BG056595 naf71f07.
	36	19	1.2	149	12	AA852172	AA852172 NHTBCae10
c	37	19	1.2	157	121	AW841413	AW841413 RC6-CN001
c	38	19	1.2	158	136	BE477554	BE477554 161314 BA
c	39	19	1.2	159	110	AW022301	AW022301 df36e10.y
	40	19	1.2	162	162	BE049235	BE049235 xw98b03.x
	41	19	1.2	164	166	BE327798	BE327798 hu28b12.x
c	42	19	1.2	166	110	AW021093	AW021093 df18f11.y
	43	19	1.2	166	148	BF448443	BF448443 7p84b05.x
	44	19	1.2	167	188	R81156	R81156 yi63d11.s3
c	45	19	1.2	168	115	AW393556	AW393556 PM3-ST030
	46	19	1.2	170	112	AW198180	AW198180 xo31f08.x
c	47	19	1.2	170	115	AW359623	AW359623 46394 MAR
c	48	19	1.2	170	115	AW359624	AW359624 46395 MAR
	49	19	1.2	171	12	AA853656	AA853656 NHTBCae07
c	50	19	1.2	172	136	BE476966	BE476966 160450 BA
c	51	19	1.2	172	136	BE477125	BE477125 160600 BA
	52	19	1.2	173	187	N92401	N92401 zb62g02.s1
	53	19	1.2	175	187	R32368	R32368 yh69f07.s1
	54	19	1.2	176	8	AA485681	AA485681 zx91c10.s
	55	19	1.2	179	12	AA853705	AA853705 NHTBCae07
	56	19	1.2	179	103	AI889746	AI889746 wo17f04.x
	57	19	1.2	179	111	AW078509	AW078509 xa99c08.x
c	58	19	1.2	181	105	AL047215	AL047215 DKFZp586H
	59	19	1.2	182	166	BE327833	BE327833 hu28g05.x
	60	19	1.2	183	112	AW197051	AW197051 xb15e12.x
	61	19	1.2	185	145	BF198150	BF198150 7p90g07.x
c	62	19	1.2	185	151	BF605964	BF605964 272886 MA
	63	19	1.2	187	111	AW080759	AW080759 xc52a09.x
	64	19	1.2	190	113	AW244128	AW244128 xa51c12.x
	65	19	1.2	191	110	AW029480	AW029480 wx12c08.x
	66	19	1.2	192	24	AI750189	AI750189 cn01a08.y
	67	19	1.2	192	103	AI888874	AI888874 wn37g07.x
	68	19	1.2	192	112	AW131080	AW131080 xf57c08.x
	69	19	1.2	193	22	AI565283	AI565283 th90d06.x
	70	19	1.2	193	103	AI935850	AI935850 wo51g04.x
	71	19	1.2	196	23	AI653927	AI653927 ty03d05.x
	72	19	1.2	196	112	AW195946	AW195946 xl85c05.x
	73	19	1.2	196	148	BF448481	BF448481 7n71b09.x
	74	19	1.2	197	22	AI613320	AI613320 tp43f12.x
	75	19	1.2	197	24	AI758599	AI758599 ty08h02.x
	76	19	1.2	197	112	AW167901	AW167901 xg55f04.x
	77	19	1.2	197	115	AW372044	AW372044 RC2-BT031
	78	19	1.2	198	21	AI537304	AI537304 tp07d08.x
	79	19	1.2	199	18	AI281578	AI281578 qu58c03.x
	80	19	1.2	199	173	BG056454	BG056454 naf69c07.
	81	19	1.2	200	104	AI971267	AI971267 wr27c07.x
	82	19	1.2	200	111	AW068347	AW068347 cn20h07.y
	83	19	1.2	201	22	AI566439	AI566439 tr95e10.x
	84	19	1.2	201	24	AI738955	AI738955 wl12d12.x
	85	19	1.2	202	19	AI360643	AI360643 qx59b04.x
	86	19	1.2	202	145	BF197773	BF197773 7p91h02.x
	87	19	1.2	203	114	AW341938	AW341938 hb73f12.x
	88	19	1.2	204	102	AI818410	AI818410 wk78e11.x
	89	19	1.2	204	111	AW068505	AW068505 cn24g02.y
	90	19	1.2	204	149	BF475728	BF475728 nac40h02.
	91	19	1.2	204	175	BG236773	BG236773 naf80b03.
	92	19	1.2	205	111	AW070647	AW070647 xa29a12.x
	93	19	1.2	205	143	BF058348	BF058348 7k22e02.x
	94	19	1.2	206	111	AW057809	AW057809 wx04b10.x
	95	19	1.2	206	112	AW149960	AW149960 xg30g06.x
	96	19	1.2	206	117	AW513616	AW513616 xo47e04.x
	97	19	1.2	206	145	BF197938	BF197938 7p87b09.x
	98	19	1.2	206	148	BF436435	BF436435 7p11g11.x
	99	19	1.2	206	173	BG059750	BG059750 naf53b05.
	100	19	1.2	206	174	BG222380	BG222380 naf75f09.
	101	19	1.2	207	12	AA852173	AA852173 NHTBCae10
	102	19	1.2	207	103	AI923684	AI923684 wn59d08.x
	103	19	1.2	207	117	AW518823	AW518823 ha43h11.x
	104	19	1.2	208	148	BF447533	BF447533 nae38e01.
	105	19	1.2	209	112	AW173442	AW173442 xj06g06.x
	106	19	1.2	212	24	AI750210	AI750210 cn01b10.x
	107	19	1.2	212	111	AW068985	AW068985 cr39g04.x
	108	19	1.2	213	103	AI922457	AI922457 wo06f07.x
	109	19	1.2	213	112	AW189452	AW189452 xl06g07.x
	110	19	1.2	214	24	AI753238	AI753238 cr07c12.x
	111	19	1.2	214	102	AI829132	AI829132 wk57f03.x
	112	19	1.2	214	104	AI982746	AI982746 wz28g01.x
	113	19	1.2	215	22	AI583570	AI583570 ts15g10.x
	114	19	1.2	215	148	BF436418	BF436418 7p11a01.x
	115	19	1.2	215	159	N67215	N67215 yz48d02.s1
	116	19	1.2	216	17	AI191457	AI191457 qe32f03.s
	117	19	1.2	216	152	BG370224	BG370224 xw89f06.x
	118	19	1.2	217	24	AI744507	AI744507 wg09a12.x
	119	19	1.2	217	24	AI753870	AI753870 cr15d01.x
	120	19	1.2	217	103	AI935815	AI935815 wo51a06.x
c	121	19	1.2	217	115	AW385796	AW385796 PM3-LT003
	122	19	1.2	218	173	BG055931	BG055931 naf83d04.
	123	19	1.2	220	111	AW080412	AW080412 xe53b08.x
	124	19	1.2	220	159	N25581	N25581 yx77b08.s1
	125	19	1.2	221	9	AA586755	AA586755 nn71c02.s
	126	19	1.2	221	158	H89108	H89108 yw25d08.s1
	127	19	1.2	222	103	AI920901	AI920901 wo15g06.x
	128	19	1.2	222	112	AW150718	AW150718 xg38c05.x
	129	19	1.2	222	114	AW273891	AW273891 xv73e12.x
	130	19	1.2	223	23	AI673819	AI673819 to73g01.x
	131	19	1.2	223	111	AW079112	AW079112 xa95h03.x
	132	19	1.2	223	112	AW192851	AW192851 xl54d06.x
	133	19	1.2	223	118	AW612927	AW612927 hh38g06.x
	134	19	1.2	224	103	AI886486	AI886486 wn39f01.x
c	135	19	1.2	224	110	AW020159	AW020159 df05d06.y
	136	19	1.2	225	21	AI537152	AI537152 tp06d06.x
c	137	19	1.2	225	144	BF094309	BF094309 CM4-UT004
	138	19	1.2	226	116	AW470186	AW470186 xw61b05.x
	139	19	1.2	227	111	AW067963	AW067963 cn22h07.x
	140	19	1.2	228	19	AI367831	AI367831 qq40c05.x

c	141	19	1.2	228	115	AW373898 AW373898 QV0-BT053
c	142	19	1.2	228	115	AW373935 AW373935 QV3-BT053
	143	19	1.2	229	15	AI074377 AI074377 oq98b08.x
	144	19	1.2	229	112	AW131138 AW131138 xf58c10.x
	145	19	1.2	229	117	AW514804 AW514804 xu88c10.x
	146	19	1.2	230	112	AW151041 AW151041 xg43h09.x
	147	19	1.2	230	115	AW393170 AW393170 CM1-ST028
	148	19	1.2	232	112	AW192315 AW192315 xl84c09.x
	149	19	1.2	232	118	AW613692 AW613692 hh77d02.x
	150	19	1.2	233	148	BF436524 BF436524 7p14e02.x
	151	19	1.2	233	163	BE068243 BE068243 MR0-BT037
	152	19	1.2	234	22	AI613434 AI613434 tp43c12.x
	153	19	1.2	238	7	AA426613 AA426613 zv47f08.s
	154	19	1.2	238	8	AA496346 AA496346 zv31d12.s
	155	19	1.2	238	18	AI281869 AI281869 qt68g04.x
	156	19	1.2	241	21	AI537464 AI537464 tp04b11.x
	157	19	1.2	242	24	AI753117 AI753117 cr05d07.x
	158	19	1.2	242	111	AW085797 AW085797 xc56g09.x
	159	19	1.2	243	20	AI446062 AI446062 tj33c08.x
	160	19	1.2	243	24	AI753880 AI753880 cr15e02.x
c	161	19	1.2	243	107	AU059177 AU059177 AU059177
	162	19	1.2	244	110	AW014698 AW014698 UI-H-BI0p
	163	19	1.2	245	159	N21969 N21969 yw30a10.s1
c	164	19	1.2	246	3	AA169264 AA169264 zp20b07.r
	165	19	1.2	247	111	AW068107 AW068107 cn23c05.x
c	166	19	1.2	248	7	AA426077 AA426077 zv47f08.r
	167	19	1.2	248	116	AW470199 AW470199 xw61d05.x
	168	19	1.2	249	162	BE049319 BE049319 xw84b02.x
	169	19	1.2	250	111	AW069811 AW069811 cr50h09.x
c	170	19	1.2	250	136	BE485222 BE485222 172173 BA
	171	19	1.2	251	159	N75080 N75080 za89h07.s1
	172	19	1.2	252	11	AA780497 AA780497 ac71a06.s
	173	19	1.2	252	24	AI754146 AI754146 cr21a01.x
	174	19	1.2	252	148	BF448500 BF448500 7n71e10.x
	175	19	1.2	253	22	AI623680 AI623680 ts22h01.x
	176	19	1.2	253	24	AI752930 AI752930 cr02d10.x
	177	19	1.2	253	24	AI754357 AI754357 cr23h01.x
	178	19	1.2	254	24	AI753100 AI753100 cr05b03.x
	179	19	1.2	254	111	AW067927 AW067927 cn22d02.x
	180	19	1.2	254	162	BE049419 BE049419 xw86a02.x
c	181	19	1.2	255	8	AA496395 AA496395 zv31d12.r
	182	19	1.2	256	24	AI750544 AI750544 cn03h07.x
	183	19	1.2	256	24	AI753408 AI753408 cr10a04.x
c	184	19	1.2	258	110	AW020567 AW020567 df11e11.y
	185	19	1.2	258	121	AW873148 AW873148 hq23e12.x
c	186	19	1.2	259	110	AW022478 AW022478 df39e12.y
	187	19	1.2	260	5	AA340382 AA340382 EST45756
	188	19	1.2	261	114	AW339169 AW339169 ha73b04.x
	189	19	1.2	264	111	AW069860 AW069860 cr52e10.x
	190	19	1.2	264	168	BF704308 BF704308 MI-P-O3-a
	191	19	1.2	265	24	AI752844 AI752844 cr01b10.x
	192	19	1.2	266	24	AI753856 AI753856 cr17e06.x
c	193	19	1.2	266	115	AW376092 AW376092 RC0-CT020
	194	19	1.2	267	9	AA621832 AA621832 nq24h10.s
	195	19	1.2	267	14	AA978078 AA978078 oq61g02.s
	196	19	1.2	269	118	AW628457 AW628457 hh95c12.x
	197	19	1.2	272	158	H88168 H88168 yw18f03.s1
	198	19	1.2	276	24	AI783883 AI783883 tr29f03.x
	199	19	1.2	278	112	AW196137 AW196137 xm05a09.x
	200	19	1.2	279	111	AW068691 AW068691 cn21h08.x
	201	19	1.2	279	112	AW193070 AW193070 xl69g12.x
	202	19	1.2	280	112	AW152521 AW152521 xg65g09.x
c	203	19	1.2	282	5	AA340068 AA340068 EST45218
	204	19	1.2	282	189	T69838 T69838 yd18g03.s1
	205	19	1.2	283	23	AI636354 AI636354 tz78c03.x
	206	19	1.2	285	112	AW194209 AW194209 xm07g10.x
	207	19	1.2	286	31	AV618814 AV618814 AV618814
	208	19	1.2	286	111	AW067892 AW067892 cn19h03.y
	209	19	1.2	287	117	AW518748 AW518748 ha39d12.x
	210	19	1.2	287	189	T50753 T50753 yb31g06.s1
	211	19	1.2	287	189	T53286 T53286 ya86h01.s1
	212	19	1.2	288	12	AA853927 AA853927 NHTBCae10
	213	19	1.2	289	3	AA167234 AA167234 zp13e10.r
	214	19	1.2	289	188	T16266 T16266 NIB1032 Nor
	215	19	1.2	292	12	AA852320 AA852320 NHTBCae12
	216	19	1.2	292	12	AA853518 AA853518 NHTBCae06
	217	19	1.2	292	18	AI274523 AI274523 qv10d03.x
	218	19	1.2	293	111	AW069526 AW069526 cr48d06.x
	219	19	1.2	293	187	R00865 R00865 ye79c01.s1
	220	19	1.2	294	12	AA853615 AA853615 NHTBCae07
	221	19	1.2	298	12	AA853294 AA853294 NHTBCae04
	222	19	1.2	298	12	AA853302 AA853302 NHTBCae04
	223	19	1.2	298	12	AA853570 AA853570 NHTBCae06
	224	19	1.2	298	12	AA853802 AA853802 NHTBCae08
	225	19	1.2	299	3	AA169128 AA169128 zp20b07.s
	226	19	1.2	299	12	AA853633 AA853633 NHTBCae07
	227	19	1.2	300	12	AA852525 AA852525 NHTBCae14
	228	19	1.2	300	12	AA852627 AA852627 NHTBCae14
	229	19	1.2	300	12	AA853104 AA853104 NHTBCae03
	230	19	1.2	300	12	AA853220 AA853220 NHTBCae04
c	231	19	1.2	300	156	C95060 C95060 C95060 Pig
	232	19	1.2	301	9	AA600069 AA600069 ag20f07.s
c	233	19	1.2	301	11	AA729258 AA729258 nx55b08.s
	234	19	1.2	301	12	AA852118 AA852118 NHTBCae10
	235	19	1.2	301	12	AA853742 AA853742 NHTBCae08
c	236	19	1.2	302	11	AA721526 AA721526 nv71a06.r
	237	19	1.2	302	12	AA853536 AA853536 NHTBCae06
	238	19	1.2	302	12	AA853620 AA853620 NHTBCae07
	239	19	1.2	302	12	AA853697 AA853697 NHTBCae07
	240	19	1.2	303	9	AA622571 AA622571 np21f05.s
	241	19	1.2	303	12	AA853348 AA853348 NHTBCae05
	242	19	1.2	303	12	AA853399 AA853399 NHTBCae05
	243	19	1.2	304	12	AA852161 AA852161 NHTBCae10
	244	19	1.2	304	12	AA852532 AA852532 NHTBCae14
	245	19	1.2	304	12	AA852904 AA852904 NHTBCae01
	246	19	1.2	304	111	AW084899 AW084899 xc65a11.x
	247	19	1.2	305	12	AA852483 AA852483 NHTBCae13
	248	19	1.2	305	12	AA853490 AA853490 NHTBCae06
	249	19	1.2	305	12	AA853583 AA853583 NHTBCae06
	250	19	1.2	305	24	AI753011 AI753011 cr03g03.x
	251	19	1.2	305	24	AI753941 AI753941 cr16d07.x
	252	19	1.2	306	12	AA852312 AA852312 NHTBCae12
c	253	19	1.2	306	159	N71173 N71173 yw30a10.r1
	254	19	1.2	307	19	AI362431 AI362431 qv92g10.x
	255	19	1.2	308	7	AA418390 AA418390 zv92e01.s
	256	19	1.2	308	12	AA853387 AA853387 NHTBCae05
	257	19	1.2	309	12	AA853601 AA853601 NHTBCae07
	258	19	1.2	309	114	AW338941 AW338941 ha68h09.x
	259	19	1.2	309	118	AW571636 AW571636 xx34d01.x
c	260	19	1.2	310	144	BF077627 BF077627 227609 MA
c	261	19	1.2	310	144	BF077698 BF077698 227692 MA
	262	19	1.2	311	12	AA852707 AA852707 NHTBCae15
	263	19	1.2	312	12	AA852896 AA852896 NHTBCae01
	264	19	1.2	312	12	AA853618 AA853618 NHTBCae07
	265	19	1.2	313	22	AI581398 AI581398 to71d12.x
	266	19	1.2	313	162	BE049583 BE049583 xw89f06.x
	267	19	1.2	314	12	AA827451 AA827451 of28e12.s
	268	19	1.2	318	20	AI452858 AI452858 tj26e04.x
	269	19	1.2	318	148	BF436512 BF436512 7p14b11.x
	270	19	1.2	318	168	BF709606 BF709606 MI-P-AY0-
	271	19	1.2	320	24	AI752838 AI752838 cr01a10.x
	272	19	1.2	320	114	AW316586 AW316586 xw09c08.x
	273	19	1.2	321	8	AA476925 AA476925 zu38c09.s
	274	19	1.2	321	9	AA564027 AA564027 nk24d04.s
	275	19	1.2	321	24	AI749049 AI749049 at38d04.x
	276	19	1.2	322	187	R26704 R26704 yh51a05.s1
c	277	19	1.2	323	111	AW067861 AW067861 cn19b01.y
	278	19	1.2	324	8	AA476901 AA476901 zu38a05.s
	279	19	1.2	324	8	AA477790 AA477790 zu34f12.s
	280	19	1.2	324	31	AV592712 AV592712 AV592712
	281	19	1.2	325	111	AW067977 AW067977 cn19b10.y
	282	19	1.2	328	3	AA174048 AA174048 zp17g03.r
c	283	19	1.2	328	5	AA340498 AA340498 EST45714
	284	19	1.2	329	111	AW084826 AW084826 xc63f04.x
	285	19	1.2	330	7	AA457435 AA457435 aa86c05.r
c	286	19	1.2	330	110	AW023052 AW023052 df47h10.y

	287	19	1.2	331	103	AI888894 AI888894 wn38b06.x
c	288	19	1.2	332	115	AW380740 AW380740 RC4-HT027
	289	19	1.2	332	188	R79240 R79240 yi88g02.s1
	290	19	1.2	333	102	AI815185 AI815185 wk72a08.x
	291	19	1.2	334	188	R65724 R65724 yi26b10.s1
	292	19	1.2	335	9	AA593946 AA593946 nn20h03.s
	293	19	1.2	335	162	BE049502 BE049502 xw87h10.x
	294	19	1.2	336	1	AA069609 AA069609 zm52c06.s
	295	19	1.2	336	24	AI754936 AI754936 cr33a02.x
	296	19	1.2	336	189	T49211 T49211 ya73c01.s1
	297	19	1.2	337	8	AA490251 AA490251 ab06h04.s
	298	19	1.2	337	24	AI750766 AI750766 cn05d06.x
c	299	19	1.2	338	7	AA457386 AA457386 aa86c05.s
	300	19	1.2	338	111	AW084942 AW084942 xc68b01.x
	301	19	1.2	339	159	N64558 N64558 yz50e09.s1
	302	19	1.2	340	112	AW190923 AW190923 xl66f06.x
	303	19	1.2	341	24	AI751708 AI751708 cn11e01.x
c	304	19	1.2	343	7	AA418439 AA418439 zv92e01.r
	305	19	1.2	343	24	AI766350 AI766350 wh60e11.x
	306	19	1.2	344	24	AI750262 AI750262 cn01e10.y
	307	19	1.2	347	22	AI623664 AI623664 ts22f06.x
	308	19	1.2	347	111	AW068075 AW068075 cn20h06.x
c	309	19	1.2	348	8	AA477014 AA477014 zu38a05.r
	310	19	1.2	348	112	AW150147 AW150147 xg48h11.x
	311	19	1.2	349	22	AI624349 AI624349 ts17b01.x
	312	19	1.2	349	24	AI750288 AI750288 cn01g04.y
	313	19	1.2	350	24	AI755247 AI755247 cr37e02.x
	314	19	1.2	350	111	AW067992 AW067992 cn19e07.y
	315	19	1.2	351	24	AI751809 AI751809 cn12a06.x
	316	19	1.2	352	23	AI702533 AI702533 tz67h07.x
	317	19	1.2	352	24	AI753687 AI753687 cr13h01.x
	318	19	1.2	352	188	R80194 R80194 yi92e03.s1
	319	19	1.2	353	31	AV593043 AV593043 AV593043
	320	19	1.2	354	24	AI750461 AI750461 cn03b09.y
	321	19	1.2	354	24	AI754295 AI754295 cr23a07.x
	322	19	1.2	355	15	AI085140 AI085140 ox95a09.s
	323	19	1.2	355	24	AI752857 AI752857 cr01d03.x
	324	19	1.2	355	24	AI755257 AI755257 cr37f02.x
	325	19	1.2	355	111	AW068297 AW068297 cn20d02.y
	326	19	1.2	356	24	AI753447 AI753447 cr10g08.x
	327	19	1.2	357	24	AI752170 AI752170 cn14b02.x
	328	19	1.2	357	24	AI752879 AI752879 cr01g01.x
	329	19	1.2	358	10	AA661579 AA661579 nu86c04.s
	330	19	1.2	358	31	AV593538 AV593538 AV593538
	331	19	1.2	359	19	AI354346 AI354346 qt78h01.x
	332	19	1.2	359	24	AI754001 AI754001 cr18d05.x
	333	19	1.2	360	3	AA218647 AA218647 zq14d04.s
c	334	19	1.2	360	10	AA661802 AA661802 nu86h12.s
	335	19	1.2	360	24	AI754861 AI754861 cr31e03.x
c	336	19	1.2	360	26	AV203409 AV203409 AV203409
c	337	19	1.2	360	107	AU097733 AU097733 AU097733
c	338	19	1.2	360	107	AU097846 AU097846 AU097846
c	339	19	1.2	360	107	AU098089 AU098089 AU098089
c	340	19	1.2	360	107	AU098095 AU098095 AU098095
c	341	19	1.2	360	107	AU098105 AU098105 AU098105
c	342	19	1.2	360	107	AU098167 AU098167 AU098167
	343	19	1.2	360	111	AW069052 AW069052 cr40f08.x
	344	19	1.2	360	158	H53594 H53594 yq87f04.s1
	345	19	1.2	361	24	AI753357 AI753357 cr09a08.x
	346	19	1.2	361	31	AV593718 AV593718 AV593718
	347	19	1.2	361	111	AW069324 AW069324 cr44c02.x
	348	19	1.2	361	169	BF818546 BF818546 MR3-CI018
	349	19	1.2	362	19	AI335520 AI335520 qt24a03.x
	350	19	1.2	362	22	AI623377 AI623377 ts18f04.x
	351	19	1.2	363	24	AI753866 AI753866 cr15c08.x
	352	19	1.2	363	111	AW069797 AW069797 cr50g02.x
	353	19	1.2	364	9	AA568353 AA568353 nl88a09.s
	354	19	1.2	364	111	AW069217 AW069217 cr41g08.x
	355	19	1.2	364	111	AW102857 AW102857 xd39c06.x
	356	19	1.2	365	24	AI754394 AI754394 cr24f01.x
	357	19	1.2	365	31	AV596030 AV596030 AV596030
c	358	19	1.2	365	31	AV596031 AV596031 AV596031
	359	19	1.2	366	24	AI753448 AI753448 cr10g09.x
	360	19	1.2	366	31	AV600140 AV600140 AV600140
c	361	19	1.2	366	110	AW023709 AW023709 df59b12.y
	362	19	1.2	366	111	AW069136 AW069136 cr42h07.x
	363	19	1.2	367	24	AI751525 AI751525 cn10e01.x
	364	19	1.2	367	116	AW469084 AW469084 hc76e10.x
	365	19	1.2	368	24	AI752213 AI752213 cn14d02.y
	366	19	1.2	368	24	AI752945 AI752945 cr02f07.x
	367	19	1.2	368	159	N67233 N67233 yz48g03.s1
	368	19	1.2	369	111	AW069575 AW069575 cr47c04.x
	369	19	1.2	370	14	AA953043 AA953043 oq70c09.s
	370	19	1.2	370	18	AI284432 AI284432 qu51c05.x
	371	19	1.2	371	4	AA255985 AA255985 zr79e03.s
	372	19	1.2	371	24	AI753297 AI753297 cr08b09.x
	373	19	1.2	371	32	AV667054 AV667054 AV667054
	374	19	1.2	372	7	AA456912 AA456912 aa90a06.s
	375	19	1.2	372	10	AA666220 AA666220 ag45c04.s
	376	19	1.2	372	24	AI752738 AI752738 cn18e01.x
	377	19	1.2	372	102	AI858855 AI858855 wl32e11.x
c	378	19	1.2	372	110	AW021107 AW021107 df18h05.y
	379	19	1.2	373	24	AI753278 AI753278 cr07h11.x
	380	19	1.2	373	24	AI753566 AI753566 cr12c10.x
	381	19	1.2	373	118	AW589566 AW589566 xo12b08.x
	382	19	1.2	374	111	AW069425 AW069425 cr45g04.x
	383	19	1.2	374	189	T99069 T99069 ye67c04.s1
	384	19	1.2	375	24	AI751585 AI751585 cn10g10.x
	385	19	1.2	375	162	BE049596 BE049596 xw89h06.x
	386	19	1.2	376	24	AI754236 AI754236 cr22b10.x
	387	19	1.2	376	173	BG055685 BG055685 naf62h07.
	388	19	1.2	377	24	AI753822 AI753822 cr17b01.x
	389	19	1.2	377	111	AW088006 AW088006 xb47c11.x
	390	19	1.2	378	111	AW068019 AW068019 cn20a06.x
	391	19	1.2	380	18	AI282426 AI282426 qu93h10.x
	392	19	1.2	381	31	AV599894 AV599894 AV599894
	393	19	1.2	381	114	AW338109 AW338109 xw66g11.x
	394	19	1.2	382	22	AI564361 AI564361 tq54h08.x
	395	19	1.2	382	188	R85592 R85592 yq28a03.s1
	396	19	1.2	383	24	AI751535 AI751535 cn10e06.x
	397	19	1.2	383	31	AV599628 AV599628 AV599628
	398	19	1.2	384	4	AA233611 AA233611 zr43c01.s
	399	19	1.2	385	7	AA428887 AA428887 zv49c01.s
	400	19	1.2	385	9	AA599475 AA599475 ag23e05.s
	401	19	1.2	385	111	AW083655 AW083655 xc49d11.x
	402	19	1.2	386	9	AA613977 AA613977 no86a09.s
	403	19	1.2	386	22	AI570836 AI570836 tr67g12.x
	404	19	1.2	386	102	AI828288 AI828288 wk82g11.x
	405	19	1.2	386	103	AI866671 AI866671 wn10a03.x
c	406	19	1.2	390	158	H89213 H89213 yw25d08.r1
	407	19	1.2	391	189	T51558 T51558 yb25f04.s1
	408	19	1.2	392	31	AV599851 AV599851 AV599851
	409	19	1.2	392	32	AV666889 AV666889 AV666889
	410	19	1.2	392	187	R35640 R35640 yh90a06.s1
	411	19	1.2	393	24	AI755081 AI755081 cr35b11.x
	412	19	1.2	393	102	AI814673 AI814673 wk66a06.x
	413	19	1.2	395	159	N72254 N72254 za01h08.s1
	414	19	1.2	397	10	AA668423 AA668423 ab84g11.s
	415	19	1.2	397	31	AV599880 AV599880 AV599880
	416	19	1.2	398	31	AV593151 AV593151 AV593151
c	417	19	1.2	398	110	AW021014 AW021014 df17g02.y
c	418	19	1.2	398	147	BF351016 BF351016 PM0-HT042
	419	19	1.2	398	159	N24817 N24817 yx98c11.s1
	420	19	1.2	398	187	R32820 R32820 yh71d02.s1
	421	19	1.2	398	189	T53376 T53376 ya88e06.s1
	422	19	1.2	399	19	AI382897 AI382897 tc20c01.x
	423	19	1.2	399	31	AV600176 AV600176 AV600176
c	424	19	1.2	400	7	AA429057 AA429057 zv49c01.r
	425	19	1.2	400	189	T56413 T56413 yb34f05.s1
	426	19	1.2	401	10	AA666230 AA666230 ag45d05.s
	427	19	1.2	401	22	AI635333 AI635333 tz80h03.x
	428	19	1.2	401	102	AI826943 AI826943 wk57b05.x
	429	19	1.2	402	7	AA457275 AA457275 aa87g08.r
	430	19	1.2	403	3	AA167585 AA167585 zp16b12.s
	431	19	1.2	403	7	AA457171 AA457171 aa88a11.r
	432	19	1.2	403	118	AW612952 AW612952 hh39d02.x

c	433	19	1.2	404	2 AA128712	AA128712 zn98b01.r
	434	19	1.2	405	10 AA678790	AA678790 ah04e06.s
	435	19	1.2	405	189 T49937	T49937 ya99e01.s1
	436	19	1.2	405	189 T55217	T55217 yb43h04.s1
	437	19	1.2	406	24 AI751732	AI751732 cn11f02.y
	438	19	1.2	406	104 AI952441	AI952441 wx73g06.x
	439	19	1.2	406	112 AW168415	AW168415 xi82h02.x
	440	19	1.2	407	22 AI589499	AI589499 tr76h06.x
	441	19	1.2	407	31 AV592588	AV592588 AV592588
	442	19	1.2	407	168 BF704165	BF704165 MI-P-O3-a
	443	19	1.2	409	31 AV593331	AV593331 AV593331
	444	19	1.2	410	162 BE046477	BE046477 hn48a09.x
	445	19	1.2	411	7 AA454820	AA454820 zx79b07.s
	446	19	1.2	411	118 AW613816	AW613816 hh74e12.x
	447	19	1.2	412	31 AV598540	AV598540 AV598540
	448	19	1.2	412	159 N67568	N67568 yz42a03.s1
	449	19	1.2	412	166 BE349988	BE349988 ht07d05.x
	450	19	1.2	413	7 AA454809	AA454809 zx79a04.s
	451	19	1.2	413	24 AI752204	AI752204 cn14c09.x
	452	19	1.2	413	24 AI752269	AI752269 cn14f08.x
	453	19	1.2	413	31 AV593651	AV593651 AV593651
	454	19	1.2	413	31 AV598120	AV598120 AV598120
	455	19	1.2	413	103 AI888745	AI888745 wn22h09.x
	456	19	1.2	415	11 AA778952	AA778952 ab69f03.r
	457	19	1.2	415	19 AI355850	AI355850 qu02d11.x
	458	19	1.2	415	31 AV592908	AV592908 AV592908
	459	19	1.2	415	31 AV598097	AV598097 AV598097
	460	19	1.2	416	19 AI338360	AI338360 qq91e02.x
	461	19	1.2	416	24 AI752177	AI752177 cn14b05.y
	462	19	1.2	416	104 AI950518	AI950518 wx51d07.x
	463	19	1.2	416	111 AW071547	AW071547 xa36d02.x
	464	19	1.2	417	3 AA176648	AA176648 zp10a01.s
	465	19	1.2	417	159 N49216	N49216 yy84f08.s1
	466	19	1.2	418	111 AW081913	AW081913 xb57b08.x
	467	19	1.2	419	31 AV592797	AV592797 AV592797
	468	19	1.2	420	10 AA666027	AA666027 ag44b10.s
	469	19	1.2	420	31 AV598091	AV598091 AV598091
	470	19	1.2	420	103 AI864605	AI864605 wg80a01.x
	471	19	1.2	421	118 AW608402	AW608402 CM0-LT006
	472	19	1.2	421	188 R69350	R69350 yi44d11.s1
	473	19	1.2	423	110 AW005970	AW005970 wz80h05.x
	474	19	1.2	423	111 AW080750	AW080750 xc51h04.x
	475	19	1.2	425	2 AA136688	AA136688 zn96d03.r
	476	19	1.2	425	10 AA670229	AA670229 ae94g08.s
	477	19	1.2	425	31 AV593415	AV593415 AV593415
	478	19	1.2	426	22 AI567094	AI567094 tq85b05.x
	479	19	1.2	426	31 AV597944	AV597944 AV597944
	480	19	1.2	426	112 AW192339	AW192339 xl44b07.x
	481	19	1.2	426	114 AW273071	AW273071 xu27f02.x
	482	19	1.2	427	31 AV597309	AV597309 AV597309
	483	19	1.2	427	31 AV599234	AV599234 AV599234
	484	19	1.2	428	12 AA838371	AA838371 of29a10.s
	485	19	1.2	428	103 AI925025	AI925025 wn27b08.x
	486	19	1.2	429	31 AV598828	AV598828 AV598828
	487	19	1.2	430	24 AI751524	AI751524 cn10d12.y
	488	19	1.2	430	158 H39824	H39824 yp01d10.s1
	489	19	1.2	431	7 AA457292	AA457292 aa88g07.s
	490	19	1.2	431	112 AW192177	AW192177 xl81a12.x
	491	19	1.2	431	188 R81155	R81155 yi63d11.s2
	492	19	1.2	432	14 AA961662	AA961662 oq80g04.s
	493	19	1.2	432	31 AV598591	AV598591 AV598591
	494	19	1.2	432	31 AV599523	AV599523 AV599523
c	495	19	1.2	433	19 AI376518	AI376518 te59a08.x
	496	19	1.2	433	24 AI751710	AI751710 cn11e02.x
	497	19	1.2	433	31 AV600035	AV600035 AV600035
	498	19	1.2	434	31 AV596424	AV596424 AV596424
	499	19	1.2	435	31 AV598184	AV598184 AV598184
	500	19	1.2	436	10 AA663136	AA663136 ab73g07.s
	501	19	1.2	436	19 AI366217	AI366217 ao87g12.x
	502	19	1.2	436	31 AV593935	AV593935 AV593935
	503	19	1.2	436	31 AV595183	AV595183 AV595183
	504	19	1.2	436	31 AV595819	AV595819 AV595819
	505	19	1.2	436	31 AV598005	AV598005 AV598005
	506	19	1.2	437	24 AI750739	AI750739 cn05c03.x
	507	19	1.2	437	24 AI751538	AI751538 cn10e08.y
	508	19	1.2	437	24 AI754060	AI754060 cr19b12.x
	509	19	1.2	437	31 AV593833	AV593833 AV593833
	510	19	1.2	437	31 AV595283	AV595283 AV595283
	511	19	1.2	437	31 AV598747	AV598747 AV598747
	512	19	1.2	438	18 AI270472	AI270472 qu88e06.x
	513	19	1.2	438	24 AI753515	AI753515 cr11f11.x
	514	19	1.2	439	2 AA121719	AA121719 zn95c10.s
	515	19	1.2	439	17 AI242508	AI242508 qu35g10.x
	516	19	1.2	439	31 AV592604	AV592604 AV592604
	517	19	1.2	439	31 AV598642	AV598642 AV598642
	518	19	1.2	439	31 AV598803	AV598803 AV598803
	519	19	1.2	439	31 AV599925	AV599925 AV599925
	520	19	1.2	439	159 N32406	N32406 yw83b02.s1
	521	19	1.2	440	2 AA121852	AA121852 zn96a09.s
	522	19	1.2	440	7 AA457605	AA457605 aa89b07.r
	523	19	1.2	440	24 AI754723	AI754723 cr29c05.x
	524	19	1.2	441	16 AI142890	AI142890 oz45a06.x
	525	19	1.2	441	19 AI367808	AI367808 qt56d10.x
	526	19	1.2	441	22 AI582556	AI582556 ts02e11.x
	527	19	1.2	442	7 AA458729	AA458729 aa87b07.r
	528	19	1.2	442	24 AI753476	AI753476 cr11b12.x
	529	19	1.2	442	112 AW168355	AW168355 xi80d09.x
	530	19	1.2	443	14 AA968534	AA968534 oq75b08.s
	531	19	1.2	443	31 AV592522	AV592522 AV592522
	532	19	1.2	443	187 N93355	N93355 zb70b08.s1
	533	19	1.2	444	31 AV593248	AV593248 AV593248
	534	19	1.2	444	32 AV667904	AV667904 AV667904
	535	19	1.2	444	164 BE221212	BE221212 hu22a05.x
	536	19	1.2	446	22 AI572735	AI572735 tr79a04.x
	537	19	1.2	446	22 AI587026	AI587026 tr55c05.x
	538	19	1.2	446	31 AV597044	AV597044 AV597044
	539	19	1.2	446	31 AV597267	AV597267 AV597267
	540	19	1.2	446	187 R50997	R50997 yg63f08.s1
	541	19	1.2	447	31 AV600105	AV600105 AV600105
	542	19	1.2	448	31 AV594091	AV594091 AV594091
	543	19	1.2	448	31 AV598408	AV598408 AV598408
	544	19	1.2	448	103 AI922973	AI922973 wn66h08.x
	545	19	1.2	449	10 AA663273	AA663273 ab80h02.s
	546	19	1.2	449	24 AI752463	AI752463 cn15g03.x
	547	19	1.2	449	31 AV598312	AV598312 AV598312
	548	19	1.2	449	111 AW079727	AW079727 xc62d10.x
	549	19	1.2	449	189 T54314	T54314 ya91c04.s3
	550	19	1.2	450	31 AV598716	AV598716 AV598716
	551	19	1.2	450	31 AV599491	AV599491 AV599491
	552	19	1.2	451	31 AV599266	AV599266 AV599266
	553	19	1.2	451	120 AW769202	AW769202 hl54f03.x
	554	19	1.2	452	2 AA128593	AA128593 zn98b01.s
	555	19	1.2	452	7 AA457381	AA457381 aa86b11.s
	556	19	1.2	452	31 AV594969	AV594969 AV594969
	557	19	1.2	452	102 AI815115	AI815115 wk71b12.x
	558	19	1.2	452	115 AW364692	AW364692 PM0-DT004
c	559	19	1.2	452	119 AW670537	AW670537 114610 MA
c	560	19	1.2	453	17 AI240057	AI240057 qh39d10.x
	561	19	1.2	453	31 AV593543	AV593543 AV593543
c	562	19	1.2	453	136 BE484555	BE484555 171246 BA
	563	19	1.2	454	11 AA759186	AA759186 ah70h07.s
	564	19	1.2	455	23 AI687267	AI687267 tp94d11.x
	565	19	1.2	455	24 AI751797	AI751797 cn11h11.x
	566	19	1.2	455	31 AV592799	AV592799 AV592799
	567	19	1.2	455	31 AV592818	AV592818 AV592818
	568	19	1.2	456	31 AV595177	AV595177 AV595177
c	569	19	1.2	456	147 BF337423	BF337423 602035148
	570	19	1.2	457	7 AA457761	AA457761 aa92c08.r
	571	19	1.2	457	10 AA664460	AA664460 ae94d10.s
	572	19	1.2	458	10 AA668540	AA668540 ag43e07.s
	573	19	1.2	458	23 AI687449	AI687449 tp96c07.x
	574	19	1.2	458	24 AI751227	AI751227 cn08e01.x
c	575	19	1.2	458	156 C27015	C27015 C27015 Rice
	576	19	1.2	459	31 AV595461	AV595461 AV595461
	577	19	1.2	459	31 AV595722	AV595722 AV595722
	578	19	1.2	460	31 AV593109	AV593109 AV593109

579	19	1.2	460	31	AV594977	AV594977 AV594977
580	19	1.2	460	31	AV614646	AV614646 AV614646
581	19	1.2	461	111	AW088830	AW088830 xd31b11.x
c 582	19	1.2	461	136	BE483186	BE483186 169258 BA
583	19	1.2	461	188	R72824	R72824 yj92b04.s1
584	19	1.2	461	190	W49666	W49666 zc41f01.s1
585	19	1.2	462	18	AI266522	AI266522 qp63f06.x
586	19	1.2	463	31	AV598657	AV598657 AV598657
c 587	19	1.2	463	232	AQ708509	AQ708509 HS_5369_A
588	19	1.2	464	10	AA663268	AA663268 ab80g05.s
589	19	1.2	464	31	AV593822	AV593822 AV593822
590	19	1.2	464	31	AV596665	AV596665 AV596665
591	19	1.2	465	188	R72990	R72990 yi38f01.r1
592	19	1.2	466	31	AV595568	AV595568 AV595568
593	19	1.2	466	31	AV596101	AV596101 AV596101
c 594	19	1.2	466	136	BE479362	BE479362 164122 BA
595	19	1.2	467	11	AA780688	AA780688 ac71g07.s
596	19	1.2	467	17	AI242707	AI242707 qt86b04.x
597	19	1.2	467	24	AI755215	AI755215 cr37a12.x
598	19	1.2	468	31	AV593288	AV593288 AV593288
599	19	1.2	468	31	AV594691	AV594691 AV594691
600	19	1.2	468	102	AI804474	AI804474 tc71f12.x
601	19	1.2	468	159	N75478	N75478 za83f01.s1
602	19	1.2	469	108	AU146031	AU146031 AU146031
603	19	1.2	470	12	AA847723	AA847723 oe20e02.s
604	19	1.2	470	14	AA985499	AA985499 oq83c03.s
605	19	1.2	471	17	AI174361	AI174361 an18h11.s
606	19	1.2	471	31	AV594463	AV594463 AV594463
607	19	1.2	471	112	AW152145	AW152145 xf75c06.x
608	19	1.2	472	22	AI635149	AI635149 tz02e11.x
609	19	1.2	472	31	AV594448	AV594448 AV594448
610	19	1.2	472	31	AV596550	AV596550 AV596550
611	19	1.2	472	31	AV597225	AV597225 AV597225
612	19	1.2	473	31	AV593392	AV593392 AV593392
613	19	1.2	474	22	AI633053	AI633053 tz01e01.x
614	19	1.2	474	31	AV593521	AV593521 AV593521
615	19	1.2	474	31	AV598038	AV598038 AV598038
616	19	1.2	474	31	AV612198	AV612198 AV612198
617	19	1.2	474	111	AW069599	AW069599 cr47e07.x
618	19	1.2	475	104	AI951166	AI951166 wx64d03.x
619	19	1.2	476	31	AV593955	AV593955 AV593955
620	19	1.2	477	15	AI079639	AI079639 ox91f11.s
621	19	1.2	478	7	AA427405	AA427405 zw32d08.s
622	19	1.2	478	10	AA665794	AA665794 ag43a03.s
623	19	1.2	478	31	AV594273	AV594273 AV594273
624	19	1.2	478	31	AV594357	AV594357 AV594357
625	19	1.2	478	31	AV595653	AV595653 AV595653
626	19	1.2	479	4	AA284822	AA284822 zt24f02.s
627	19	1.2	479	8	AA477812	AA477812 zu38h03.s
628	19	1.2	479	20	AI431945	AI431945 th32e10.x
629	19	1.2	479	31	AV595936	AV595936 AV595936
630	19	1.2	481	9	AA593813	AA593813 nn17a04.s
631	19	1.2	481	24	AI753591	AI753591 cr12f06.x
632	19	1.2	481	31	AV595798	AV595798 AV595798
633	19	1.2	482	7	AA457264	AA457264 aa91f08.s
634	19	1.2	482	9	AA613975	AA613975 no86a07.s
635	19	1.2	482	10	AA668442	AA668442 ab85e07.s
636	19	1.2	482	10	AA668635	AA668635 ag45h06.s
637	19	1.2	483	17	AI185087	AI185087 qe46h09.x
638	19	1.2	483	31	AV593365	AV593365 AV593365
639	19	1.2	483	111	AW069105	AW069105 cr41d02.x
640	19	1.2	484	111	AW069601	AW069601 cr47e09.x
c 641	19	1.2	484	139	BE721467	BE721467 188652 MA
642	19	1.2	485	24	AI755173	AI755173 cr36e05.x
643	19	1.2	485	31	AV593749	AV593749 AV593749
644	19	1.2	486	9	AA580405	AA580405 nn13h02.s
645	19	1.2	486	112	AW198050	AW198050 xo27h03.x
646	19	1.2	487	111	AW067792	AW067792 cn19c04.x
647	19	1.2	487	111	AW068870	AW068870 cn25h02.x
648	19	1.2	488	10	AA668601	AA668601 ag45e02.s
649	19	1.2	488	14	AA977821	AA977821 oq78a11.s
650	19	1.2	488	116	AW473961	AW473961 hb12a11.x
651	19	1.2	490	9	AA594770	AA594770 nn86b09.s
652	19	1.2	490	22	AI582561	AI582561 ts02f04.x
653	19	1.2	491	9	AA600282	AA600282 af01f12.s
c 654	19	1.2	491	136	BE476709	BE476709 160086 BA
655	19	1.2	492	9	AA600269	AA600269 af01e10.s
656	19	1.2	492	9	AA600271	AA600271 af01e12.s
657	19	1.2	494	12	AA827555	AA827555 of33a10.s
658	19	1.2	494	22	AI581422	AI581422 to71g08.x
659	19	1.2	495	7	AA446927	AA446927 zw85b04.s
660	19	1.2	495	14	AA983262	AA983262 oq78e09.s
661	19	1.2	495	103	AI888567	AI888567 wn33a10.x
662	19	1.2	495	116	AW475081	AW475081 hb02g02.x
663	19	1.2	496	31	AV594431	AV594431 AV594431
664	19	1.2	497	104	AI983016	AI983016 wt46g12.x
665	19	1.2	498	114	AW337149	AW337149 xw81e07.x
666	19	1.2	499	9	AA600020	AA600020 ag20a06.s
667	19	1.2	499	11	AA772878	AA772878 ab70d01.r
668	19	1.2	500	104	AI982991	AI982991 wt46d11.x
c 669	19	1.2	502	138	BE663181	BE663181 143867 MA
670	19	1.2	503	9	AA600032	AA600032 ag20b08.s
671	19	1.2	504	22	AI590419	AI590419 ts10c12.x
672	19	1.2	504	24	AI754676	AI754676 cr28f05.x
673	19	1.2	505	20	AI445593	AI445593 tj26b10.x
674	19	1.2	506	7	AA453760	AA453760 aa19c03.s
675	19	1.2	506	21	AI520803	AI520803 ti07g10.x
676	19	1.2	508	20	AI433929	AI433929 ti11g02.x
c 677	19	1.2	508	110	AW020015	AW020015 df02h05.y
678	19	1.2	509	14	AA988677	AA988677 oq88f12.s
679	19	1.2	509	21	AI537514	AI537514 tp04e03.x
680	19	1.2	509	108	AU145887	AU145887 AU145887
681	19	1.2	511	14	AA953103	AA953103 oq70f06.s
c 682	19	1.2	512	3	AA181457	AA181457 zp16d05.r
683	19	1.2	513	7	AA453741	AA453741 aa19a03.s
684	19	1.2	513	9	AA602898	AA602898 np39d04.s
685	19	1.2	513	31	AV602865	AV602865 AV602865
c 686	19	1.2	514	10	AA653779	AA653779 ns90d07.s
687	19	1.2	514	20	AI436455	AI436455 ti09c11.x
688	19	1.2	514	102	AI802219	AI802219 tj36b11.x
689	19	1.2	514	112	AW191918	AW191918 xl77b09.x
690	19	1.2	515	7	AA405497	AA405497 zw36d05.s
691	19	1.2	515	22	AI571460	AI571460 tr56d09.x
692	19	1.2	517	9	AA599355	AA599355 ag35h07.s
693	19	1.2	517	11	AA780459	AA780459 ac65b07.s
694	19	1.2	517	11	AA780725	AA780725 ac68a06.s
c 695	19	1.2	517	233	AQ756735	AQ756735 HS_5407_A
696	19	1.2	518	111	AW068554	AW068554 cn21b03.x
697	19	1.2	521	3	AA179999	AA179999 zp16d04.s
698	19	1.2	522	14	AA987731	AA987731 oq94c08.s
699	19	1.2	522	22	AI567881	AI567881 tr63a06.x
700	19	1.2	523	8	AA486114	AA486114 ab14f08.s
701	19	1.2	523	22	AI631341	AI631341 tz83b09.x
702	19	1.2	523	104	AI953832	AI953832 wx70a05.x
703	19	1.2	524	13	AA937318	AA937318 oh35e09.s
704	19	1.2	525	18	AI284036	AI284036 qt72c05.x
705	19	1.2	525	22	AI589473	AI589473 tr76e07.x
706	19	1.2	526	14	AA975668	AA975668 oq63f12.s
707	19	1.2	526	16	AI128311	AI128311 qc60f02.x
708	19	1.2	526	18	AI287470	AI287470 qv81e09.x
709	19	1.2	526	22	AI571470	AI571470 tr56e08.x
710	19	1.2	526	22	AI621204	AI621204 tu88g11.x
c 711	19	1.2	526	31	AV594432	AV594432 AV594432
712	19	1.2	526	111	AW081909	AW081909 xb57b01.x
c 713	19	1.2	526	173	BG110545	BG110545 602278974
714	19	1.2	527	14	AA987722	AA987722 oq94b01.s
715	19	1.2	527	102	AI828211	AI828211 wk81e11.x
c 716	19	1.2	527	165	BE236174	BE236174 143868 MA
717	19	1.2	530	9	AA599996	AA599996 ag29g04.s
718	19	1.2	531	9	AA617730	AA617730 np51a09.s
719	19	1.2	531	18	AI281680	AI281680 qu79c12.x
720	19	1.2	531	20	AI473655	AI473655 tm02f08.x
721	19	1.2	532	159	N32090	N32090 yw97f07.s1
722	19	1.2	533	159	N26572	N26572 yx91b07.s1
723	19	1.2	534	17	AI188285	AI188285 qd07e03.x
724	19	1.2	535	9	AA551267	AA551267 nj35g06.s

	725	19	1.2	535	103	AI922136 AI922136 wn88h04.x
	726	19	1.2	536	8	AA496278 AA496278 aa94d08.r
	727	19	1.2	537	4	AA293101 AA293101 zt55c01.s
	728	19	1.2	537	112	AW150699 AW150699 xg38a03.x
	729	19	1.2	537	114	AW276582 AW276582 xr16b11.x
	730	19	1.2	540	23	AI708458 AI708458 as97e09.x
	731	19	1.2	540	31	AV612207 AV612207
c	732	19	1.2	541	4	AA284550 AA284550 zt24f02.r
	733	19	1.2	543	9	AA570534 AA570534 nk65e07.s
	734	19	1.2	544	4	AA256158 AA256158 zr79b02.s
	735	19	1.2	549	22	AI590390 AI590390 ts09h07.x
	736	19	1.2	550	118	AW602027 AW602027 QV3-BT053
	737	19	1.2	551	4	AA292300 AA292300 zt51e08.s
c	738	19	1.2	552	4	AA256247 AA256247 zr79e03.r
	739	19	1.2	552	8	AA490138 AA490138 ab05f03.s
c	740	19	1.2	552	31	AV593834 AV593834
	741	19	1.2	552	103	AI921143 AI921143 wo21a12.x
c	742	19	1.2	552	110	AW022245 AW022245 df34h05.y
	743	19	1.2	554	9	AA599428 AA599428 ag23d03.s
	744	19	1.2	554	20	AI453747 AI453747 tj39g10.x
	745	19	1.2	555	22	AI567549 AI567549 tr89g08.x
c	746	19	1.2	555	141	BE845745 BE845745 232890 BA
	747	19	1.2	559	8	AA487276 AA487276 aa94e11.s
	748	19	1.2	559	9	AA599762 AA599762 ag33a02.s
	749	19	1.2	559	9	AA599979 AA599979 ag28g08.s
	750	19	1.2	560	9	AA586696 AA586696 nn42h07.s
	751	19	1.2	560	31	AV595687 AV595687
c	752	19	1.2	560	165	BE236099 BE236099 143776 MA
c	753	19	1.2	561	31	AV599926 AV599926
	754	19	1.2	562	9	AA599521 AA599521 ag26f02.s
	755	19	1.2	562	31	AV599702 AV599702
	756	19	1.2	564	23	AI674910 AI674910 wc73f07.x
	757	19	1.2	565	9	AA564970 AA564970 nj24e04.s
c	758	19	1.2	565	228	AQ415712 AQ415712 RPCI-11-2
	759	19	1.2	566	22	AI567509 AI567509 tr89b11.x
	760	19	1.2	566	108	AU144820 AU144820 AU144820
	761	19	1.2	567	9	AA599483 AA599483 ag23f01.s
	762	19	1.2	568	7	AA456909 AA456909 aa90a02.s
	763	19	1.2	568	8	AA487514 AA487514 aa95d04.s
c	764	19	1.2	569	7	AA456983 AA456983 aa90a02.r
	765	19	1.2	570	4	AA293155 AA293155 zt55d08.s
	766	19	1.2	574	8	AA489810 AA489810 ab01h03.s
	767	19	1.2	575	103	AI925509 AI925509 wo31d06.x
c	768	19	1.2	578	31	AV598541 AV598541
	769	19	1.2	579	20	AI446391 AI446391 tj10e11.x
	770	19	1.2	579	23	AI694179 AI694179 wc83g07.x
	771	19	1.2	580	9	AA599208 AA599208 ag34b10.s
	772	19	1.2	580	102	AI811250 AI811250 tw72h08.x
	773	19	1.2	581	112	AW190727 AW190727 x141g02.x
	774	19	1.2	583	9	AA599300 AA599300 ag35c12.s
	775	19	1.2	583	20	AI453365 AI453365 tj37c09.x
	776	19	1.2	584	102	AI828570 AI828570 wl42d04.x
	777	19	1.2	586	117	AW514823 AW514823 xu88h08.x
	778	19	1.2	588	103	AI912433 AI912433 tz05e03.x
	779	19	1.2	589	18	AI281603 AI281603 qu53a04.x
	780	19	1.2	590	8	AA489802 AA489802 ab01g05.s
	781	19	1.2	590	108	AU144879 AU144879 AU144879
	782	19	1.2	593	112	AW192764 AW192764 x150h05.x
c	783	19	1.2	594	141	BE846165 BE846165 232194 BA
	784	19	1.2	601	9	AA564675 AA564675 nj22e09.s
	785	19	1.2	601	31	AV596154 AV596154
	786	19	1.2	601	121	AW857365 AW857365 CM3-CT031
	787	19	1.2	603	115	AW373802 AW373802 QV3-BT053
	788	19	1.2	605	103	AI906142 AI906142 CM-BT106-
	789	19	1.2	605	111	AW077898 AW077898 fj67h06.x
	790	19	1.2	610	104	AI963411 AI963411 wt62h03.x
	791	19	1.2	611	162	BE049316 BE049316 xw84a09.x
	792	19	1.2	612	103	AI922681 AI922681 wo11d01.x
	793	19	1.2	612	249	AZ790576 AZ790576 2M0039H12
	794	19	1.2	618	114	AW272699 AW272699 xu21h10.x
	795	19	1.2	619	9	AA595560 AA595560 ni47c08.s
	796	19	1.2	621	111	AW082115 AW082115 xb60g01.x
	797	19	1.2	622	22	AI587173 AI587173 tr54c07.x
	798	19	1.2	624	23	AI675919 AI675919 wc73b04.x
	799	19	1.2	624	24	AI752736 AI752736 cn18d12.x
	800	19	1.2	625	105	AL035880 AL035880 DKFZp564C
	801	19	1.2	626	173	BG109126 BG109126 602281319
	802	19	1.2	628	10	AA669785 AA669785 ag36a11.s
c	803	19	1.2	629	115	AW373808 AW373808 QV3-BT053
	804	19	1.2	629	118	AW577398 AW577398 QV3-BT053
c	805	19	1.2	630	118	AW577403 AW577403 QV3-BT053
	806	19	1.2	633	9	AA594126 AA594126 nm82b12.s
	807	19	1.2	634	104	AI963366 AI963366 wt62c07.x
	808	19	1.2	639	20	AI445562 AI445562 tj30g02.x
	809	19	1.2	640	10	AA669422 AA669422 ac17c02.s
	810	19	1.2	641	17	AI185516 AI185516 qe54f11.x
	811	19	1.2	641	115	AW373788 AW373788 QV3-BT053
	812	19	1.2	641	162	BE049241 BE049241 xw98c07.x
	813	19	1.2	642	10	AA669998 AA669998 ag38h04.s
c	814	19	1.2	642	115	AW373902 AW373902 QV0-BT053
c	815	19	1.2	642	115	AW373939 AW373939 QV3-BT053
c	816	19	1.2	643	118	AW577418 AW577418 QV3-BT053
	817	19	1.2	644	24	AI745640 AI745640 tr24b02.x
	818	19	1.2	645	23	AI685149 AI685149 wa75b03.x
	819	19	1.2	645	102	AI801622 AI801622 to91h01.x
c	820	19	1.2	646	30	AV406305 AV406305 AV406305
c	821	19	1.2	646	105	AL047865 AL047865 DKFZp586M
	822	19	1.2	647	8	AA522675 AA522675 ni39b01.s
	823	19	1.2	654	104	AI963017 AI963017 wt26b10.x
	824	19	1.2	655	112	AW152444 AW152444 xg64g01.x
	825	19	1.2	656	115	AW373787 AW373787 QV3-BT053
	826	19	1.2	660	10	AA669973 AA669973 ag38f02.s
	827	19	1.2	663	121	AW857368 AW857368 CM3-CT031
	828	19	1.2	666	23	AI677812 AI677812 wc80f09.x
	829	19	1.2	667	8	AA522696 AA522696 ni39d01.s
	830	19	1.2	668	17	AI227911 AI227911 EST224606
	831	19	1.2	672	102	AI796681 AI796681 wh59e02.x
	832	19	1.2	672	103	AI925580 AI925580 wo33h02.x
	833	19	1.2	675	23	AI696990 AI696990 wc76f02.x
	834	19	1.2	677	121	AW860519 AW860519 QV0-CT038
	835	19	1.2	682	22	AI624675 AI624675 ts43e02.x
	836	19	1.2	688	23	AI676195 AI676195 wc72b07.x
	837	19	1.2	689	22	AI571571 AI571571 tr85g11.x
	838	19	1.2	690	22	AI570780 AI570780 tr67a10.x
	839	19	1.2	736	9	AA554805 AA554805 ni34a04.s
	840	19	1.2	738	23	AI684315 AI684315 tw97g03.x
	841	19	1.2	738	111	AW084616 AW084616 xa45d01.x
	842	19	1.2	748	22	AI567787 AI567787 tr62f01.x
	843	19	1.2	771	12	AA788961 AA788961 ag40d11.s
	844	19	1.2	801	22	AI624300 AI624300 ts28h08.x
c	845	19	1.2	806	226	AQ287651 AQ287651 nbxb0031C
	846	19	1.2	823	12	AA789233 AA789233 aj27e06.s
	847	19	1.2	825	102	AI829366 AI829366 wk64b07.x
c	848	19	1.2	829	107	AU117762 AU117762 AU117762
c	849	19	1.2	840	30	AV405942 AV405942
	850	19	1.2	851	168	BF700296 BF700296 602127720
c	851	19	1.2	853	172	BG026188 BG026188 602291521
	852	19	1.2	866	10	AA703999 AA703999 ag46h05.s
	853	19	1.2	869	10	AA704140 AA704140 ag47g04.s
	854	19	1.2	911	144	BF098988 BF098988 601750494
	855	19	1.2	934	106	AL568796 AL568796
	856	19	1.2	995	222	CNS05GN7 AL336508 Tetraodon
c	857	19	1.2	1028	222	CNS04XV8 AL312173 Tetraodon
c	858	19	1.2	1034	173	BG111131 BG111131 602281538
	859	19	1.2	1053	23	AI654385 AI654385 tq90f02.x
c	860	19	1.2	1073	174	BG113870 BG113870 602284322
	861	18	1.2	88	246	AZ584971 1M0390C03
c	862	18	1.2	97	14	AA993429 AA993429 ot96d11.s
	863	18	1.2	116	12	AA821411 AA821411 vs67h04.r
c	864	18	1.2	146	166	BE336328 BE336328 601087094
	865	18	1.2	163	148	BF417740 BF417740 UI-R-CN0-
	866	18	1.2	164	152	BG373163 BG373163 UI-R-CV0-
c	867	18	1.2	166	12	AA840346 AA840346 vw92a07.r
c	868	18	1.2	166	32	AV671832 AV671832 AV671832
	869	18	1.2	167	145	BF150668 BF150668 uy87e12.x
c	870	18	1.2	186	154	BG518894 BG518894 602578194

c 871	18	1.2	195	108	AV015513	AV015513 AV015513
872	18	1.2	202	145	BF150722	BF150722 uy87h10.x
c 873	18	1.2	208	144	BF118205	BF118205 uz13h02.y
874	18	1.2	213	144	BF118355	BF118355 uy94c08.x
875	18	1.2	214	146	BF224719	BF224719 uz03d05.x
c 876	18	1.2	214	165	BE284823	BE284823 601091138
877	18	1.2	216	146	BF227015	BF227015 uz35e12.x
878	18	1.2	216	146	BF227682	BF227682 uz32e09.x
879	18	1.2	217	107	AU058976	AU058976 AU058976
880	18	1.2	221	102	AI841667	AI841667 UI-M-AO0-
c 881	18	1.2	222	4	AA240871	AA240871 mw17a06.r
c 882	18	1.2	222	188	T01463	T01463 wEST02184 E
c 883	18	1.2	224	25	AV150597	AV150597 AV150597
c 884	18	1.2	225	135	BB484713	BB484713 BB484713
885	18	1.2	227	146	BF224552	BF224552 uz07g08.x
886	18	1.2	228	173	BG090379	BG090379 mab96d02.
887	18	1.2	231	146	BF225013	BF225013 uz13h02.x
888	18	1.2	232	18	AI325244	AI325244 ms34h08.x
c 889	18	1.2	236	27	AV245242	AV245242 AV245242
890	18	1.2	236	151	BF661619	BF661619 maa63f10.
c 891	18	1.2	237	27	AV245104	AV245104 AV245104
c 892	18	1.2	237	30	AV378933	AV378933 AV378933
893	18	1.2	238	113	AW208563	AW208563 uo61c06.x
c 894	18	1.2	239	29	AV322242	AV322242 AV322242
c 895	18	1.2	246	108	AV007958	AV007958 AV007958
c 896	18	1.2	247	132	BB372788	BB372788 BB372788
c 897	18	1.2	248	109	AV089573	AV089573 AV089573
c 898	18	1.2	248	160	BB540477	BB540477 BB540477
899	18	1.2	249	21	AI555384	AI555384 UI-R-C2p-
c 900	18	1.2	249	26	AV209786	AV209786 AV209786
901	18	1.2	249	28	AV280329	AV280329 AV280329
c 902	18	1.2	249	135	BB485433	BB485433 BB485433
c 903	18	1.2	253	135	BB485397	BB485397 BB485397
c 904	18	1.2	255	18	AI322257	AI322257 ms34h08.y
905	18	1.2	256	146	BF226836	BF226836 uz16c08.x
c 906	18	1.2	257	3	AA168805	AA168805 ms34h08.r
c 907	18	1.2	258	134	BB435952	BB435952 BB435952
c 908	18	1.2	258	161	BB599432	BB599432 BB599432
c 909	18	1.2	260	4	AA265649	AA265649 mw14d09.r
c 910	18	1.2	264	169	BF748807	BF748807 MR0-BN011
c 911	18	1.2	268	132	BB373855	BB373855 BB373855
c 912	18	1.2	270	25	AV165093	AV165093 AV165093
913	18	1.2	271	173	BG060649	BG060649 L0906B11-
914	18	1.2	273	103	AI865239	AI865239 wk08h11.x
c 915	18	1.2	274	29	AV350926	AV350926 AV350926
c 916	18	1.2	274	132	BB376924	BB376924 BB376924
c 917	18	1.2	276	136	BE522366	BE522366 M25D3XTM
918	18	1.2	277	111	AW069093	AW069093 cr41b11.x
c 919	18	1.2	281	26	AV227153	AV227153 AV227153
920	18	1.2	285	27	AV237624	AV237624 AV237624
c 921	18	1.2	286	123	BB007065	BB007065 BB007065
922	18	1.2	286	163	BE128206	BE128206 DEPA1955
923	18	1.2	287	117	AW532953	AW532953 UI-R-BU0-
c 924	18	1.2	288	8	AA498114	AA498114 vh33e06.r
925	18	1.2	288	25	AV149603	AV149603 AV149603
c 926	18	1.2	289	123	BB009908	BB009908 BB009908
c 927	18	1.2	289	125	BB091072	BB091072 BB091072
928	18	1.2	289	145	BF148534	BF148534 uy82e03.x
c 929	18	1.2	292	132	BB371647	BB371647 BB371647
c 930	18	1.2	292	160	BB538720	BB538720 BB538720
c 931	18	1.2	294	26	AV226457	AV226457 AV226457
932	18	1.2	298	26	AV227451	AV227451 AV227451
c 933	18	1.2	300	132	BB379488	BB379488 BB379488
934	18	1.2	300	145	BF148434	BF148434 uy81a11.x
c 935	18	1.2	305	25	AV171430	AV171430 AV171430
c 936	18	1.2	306	25	AV138339	AV138339 AV138339
c 937	18	1.2	308	4	AA230929	AA230929 mw08d12.r
938	18	1.2	308	109	AV046787	AV046787 AV046787
c 939	18	1.2	308	132	BB385371	BB385371 BB385371
940	18	1.2	308	146	BF226830	BF226830 uz16c02.x
c 941	18	1.2	309	132	BB376752	BB376752 BB376752
942	18	1.2	310	165	BE225879	BE225879 ia20h11.x
c 943	18	1.2	310	189	W11321	W11321 ma78c07.r1
944	18	1.2	311	20	AI413740	AI413740 me84e04.x
c 945	18	1.2	315	129	BB251873	BB251873 BB251873
946	18	1.2	318	146	BF226746	BF226746 uz15d10.x
c 947	18	1.2	318	188	T01462	T01462 wEST02183 E
c 948	18	1.2	318	190	W16311	W16311 mb55f06.r1
949	18	1.2	324	105	AL036194	AL036194 DKFZp564G
950	18	1.2	324	116	AW434444	AW434444 UI-R-BJ0p
951	18	1.2	325	238	AZ082311	AZ082311 RPCI-23-1
c 952	18	1.2	328	1	AA050608	AA050608 mj19h04.r
c 953	18	1.2	329	247	AZ649832	AZ649832 1M0519J10
c 954	18	1.2	330	12	AA791164	AA791164 vv91f11.r
c 955	18	1.2	330	20	AI425767	AI425767 me84e04.y
956	18	1.2	330	249	AZ769004	AZ769004 1M0569H14
c 957	18	1.2	337	117	AW530284	AW530284 UI-R-BU0-
c 958	18	1.2	338	132	BB377752	BB377752 BB377752
c 959	18	1.2	343	117	AW506372	AW506372 GE1631 Gl
960	18	1.2	343	163	BE116168	BE116168 UI-R-BS1-
c 961	18	1.2	344	12	AA799161	AA799161 vn41h05.r
962	18	1.2	344	165	BE225955	BE225955 ia22d07.x
c 963	18	1.2	345	15	AI045586	AI045586 UI-R-C1-k
c 964	18	1.2	345	149	BF452923	BF452923 maa63f10.
c 965	18	1.2	348	3	AA175736	AA175736 ms95c04.r
966	18	1.2	351	21	AI500865	AI500865 UI-R-C2p-
967	18	1.2	351	116	AW433712	AW433712 UI-R-BJ0p
c 968	18	1.2	353	1	AA008230	AA008230 mg70e11.r
c 969	18	1.2	353	190	W64655	W64655 md70d01.r1
970	18	1.2	357	256	B70470	B70470 CIT-HSP-205
c 971	18	1.2	360	26	AV193532	AV193532 AV193532
c 972	18	1.2	360	156	C60359	C60359 C60359 Yuji
c 973	18	1.2	363	3	AA146453	AA146453 mr05c03.r
c 974	18	1.2	363	189	W07999	W07999 mb43a05.r1
975	18	1.2	364	23	AI706254	AI706254 UI-R-AC0-
976	18	1.2	364	113	AW228102	AW228102 up13d01.x
c 977	18	1.2	366	190	W41217	W41217 mc42b06.r1
978	18	1.2	367	14	AA992231	AA992231 ot29e05.s
c 979	18	1.2	368	150	BF545674	BF545674 UI-R-C2p-
980	18	1.2	369	16	AI137107	AI137107 UI-R-C2p-
c 981	18	1.2	370	4	AA274801	AA274801 va92a10.r
982	18	1.2	371	102	AI838738	AI838738 UI-M-AO0-
983	18	1.2	372	11	AA756734	AA756734 vv65a05.r
984	18	1.2	373	153	BG412243	BG412243 OV2_39_F0
c 985	18	1.2	373	224	AQ103689	AQ103689 HS_3092_B
986	18	1.2	381	138	BE690753	BE690753 uw58f02.x
c 987	18	1.2	387	174	BG176387	BG176387 602338052
c 988	18	1.2	388	17	AI232299	AI232299 EST228987
c 989	18	1.2	388	239	AZ140167	AZ140167 SP_0005_A
990	18	1.2	391	113	AW259634	AW259634 up31c12.x
c 991	18	1.2	391	150	BF528406	BF528406 602043756
992	18	1.2	393	111	AW123300	AW123300 UI-M-BH2.
993	18	1.2	394	113	AW240865	AW240865 uq37a08.x
994	18	1.2	394	115	AW412757	AW412757 uq48h06.x
995	18	1.2	398	24	AI753333	AI753333 cr08f08.x
996	18	1.2	398	142	BE956428	BE956428 UI-M-BH4-
997	18	1.2	400	244	AZ482810	AZ482810 1M0308L05
998	18	1.2	401	20	AI481729	AI481729 vh33e06.x
c 999	18	1.2	401	144	BF123989	BF123989 601759675
c1000	18	1.2	402	4	AA230997	AA230997 mw33a06.r

ALIGNMENTS

RESULT 1
AQ026293
LOCUS AQ026293 559 bp DNA GSS 30-JUN-1998
DEFINITION l(3)j1C7 Drosophila melanogaster P lethal line Drosophila melanogaster genomic Sequence recovered from Both 5' and 3' ends of P element, DNA sequence.
ACCESSION AQ026293
VERSION AQ026293.1 GI:3266578
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster

```
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 559)
  AUTHORS   Spradling,A.C., Stern,D., Beaton,A., Rehm,E.J., Laverty,T., Mozden
            ,N., Misra,S. and Rubin,G.M.
  TITLE     The BDGP gene disruption project: Single P element insertions
            mutating 30% of Drosophila autosomal genes
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.

            Sequence orientation is forward strand relative to 5' end of P
            element

            The P element insertion position is base 552 in the 559 bases. This
            insertion position refers to the first base of the 8 base target
            recognition sequence.
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster P lethal line"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains a single P
                     transposable element insertion that is thought to cause
                     either lethality or sterility. The resultant fragment for
                     each strain was directly sequenced to determine the
                     genomic sequence at the site of insertion. Details of the
                     protocols used can be found at
                     http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."
BASE COUNT      180 a    113 c    114 g    152 t
ORIGIN


  Query Match             2.0%;  Score 31;  DB 223;  Length 559;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-06;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1506 gcgcagcacatggactctaatgatggattct 1536
           |||||||||||||||||||||||||||||||
Db       1 GCGCAGCACATGGACTCTAATGATGGATTCT 31


RESULT   2
BE781852/c
LOCUS       BE781852      904 bp    mRNA            EST       20-OCT-2000
DEFINITION  601470173F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3872946 5',
            mRNA sequence.
ACCESSION   BE781852
VERSION     BE781852.1  GI:10203050
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 904)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9628  row: b  column: 19
            High quality sequence stop: 580.
FEATURES             Location/Qualifiers
     source          1. .904
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3872946"
                     /clone_lib="NIH_MGC_67"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      267 a    147 c    193 g    297 t
ORIGIN


  Query Match             1.4%;  Score 22;  DB 140;  Length 904;
  Best Local Similarity   100.0%;  Pred. No. 0.81;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     348 ctatatatacttcacatacatg 369
           |||||||||||||||||||||
Db     461 CTATATATACTTCACATACATG 440


RESULT   3
AW176349/c
LOCUS       AW176349      104 bp    mRNA            EST       16-NOV-1999
DEFINITION  QV0-BT0233-010999-010-e12 BT0233 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW176349
VERSION     AW176349.1  GI:6442386
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 104)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV0&t2=QV0-BT0233-
            010999-010-e12&t3=1999-09-01&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 24
            High quality sequence stop: 104.
FEATURES             Location/Qualifiers
     source          1. .104
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BT0233"
                     /dev_stage="Adult"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT       28 a     24 c     15 g     37 t
```

```
ORIGIN


  Query Match             1.4%;  Score 21;  DB 112;  Length 104;
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    605 attttccaaagtccatgtgga 625
          ||||||||||||||||||||
Db     22 ATTTTCCAAAGTCCATGTGGA 2


RESULT   4
AI752336
LOCUS       AI752336      229 bp    mRNA            EST       22-JUN-1999
DEFINITION  cn15a10.x1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
            clone NHTBC_cn15a10 random, mRNA sequence.
ACCESSION   AI752336
VERSION     AI752336.1  GI:5130600
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 229)
  AUTHORS   Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
            Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M., Robey
            ,P.G., Hotchkiss,R.N. and Francomano,C.A.
  TITLE     SGAP: The Skeletal Genome Anatomy Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Libin Jia
            Medical Genetics Branch
            National Human Genome Research Institute
            10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
            Tel: 301-402-4877
            Fax: 301-496-7157
            Email: libin@helix.nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
            Plate: 15  row: a  column: 10
            Seq primer: -21M13 forward primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .229
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NHTBC_cn15a10"
                     /clone_lib="Normal Human Trabecular Bone Cells"
                     /sex="Female"
                     /tissue_type="Bone"
                     /cell_type="Trabecular Bone Cells"
                     /lab_host="SURE"
                     /note="Organ: Hip; Vector: pBluescript; Site_1: EcoRI"
BASE COUNT       50 a     27 c     81 g     69 t      2 others
ORIGIN


  Query Match             1.4%;  Score 21;  DB 24;  Length 229;
  Best Local Similarity   100.0%;  Pred. No. 2.9;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    605 attttccaaagtccatgtgga 625
          ||||||||||||||||||||
Db     95 ATTTTCCAAAGTCCATGTGGA 115


RESULT   5
AA457068
LOCUS       AA457068      313 bp    mRNA            EST       06-JUN-1997
DEFINITION  aa90f05.r1 Stratagene fetal retina 937202 Homo sapiens cDNA clone
            IMAGE:838593 5' similar to gb:K01228 PROCOLLAGEN ALPHA 1(I) CHAIN
            PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   AA457068
VERSION     AA457068.1  GI:2179788
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 313)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .313
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:838593"
                     /clone_lib="Stratagene fetal retina 937202"
                     /sex="mixed"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. Pooled
                     retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
                     Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       78 a     53 c     95 g     87 t
ORIGIN


  Query Match             1.4%;  Score 21;  DB 7;  Length 313;
  Best Local Similarity   100.0%;  Pred. No. 2.9;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    605 attttccaaagtccatgtgga 625
          ||||||||||||||||||||
Db     84 ATTTTCCAAAGTCCATGTGGA 104


RESULT   6
BE357901
LOCUS       BE357901      590 bp    mRNA            EST       20-JUL-2000
DEFINITION  DG1_23_H02.b1_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   BE357901
VERSION     BE357901.1  GI:9299458
KEYWORDS    EST.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
            ,L.H.
  TITLE     An EST database from Sorghum: dark-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
```

```
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: JEN REV
            High quality sequence stop: 590
            POLYA=No.
FEATURES              Location/Qualifiers
     source           1. .590
                      /organism="Sorghum bicolor"
                      /db_xref="taxon:4558"
                      /clone_lib="Dark Grown 1 (DG1)"
                      /note="Organ: 5-day-old dark-grown seedlings; Vector:
                      Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
                      made from poly-A RNA in the cloning vector lambda ZAP II.
                      Clones to be sequenced were prepared by mass excision."
BASE COUNT      135 a    157 c    150 g    148 t
ORIGIN


  Query Match             1.4%;  Score 21;  DB 166;  Length 590;
  Best Local Similarity   100.0%;  Pred. No. 3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     20 ttttctgtacagttctgcttt 40
          |||||||||||||||||||||
Db    252 TTTTCTGTACAGTTCTGCTTT 272


RESULT   7
BF574583
LOCUS       BF574583      820 bp    mRNA            EST       12-DEC-2000
DEFINITION  602133237F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4288232 5',
            mRNA sequence.
ACCESSION   BF574583
VERSION     BF574583.1  GI:11648295
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 820)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
            cDNA Library Preparation: CLONETECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1127  row: j  column: 09
            High quality sequence stop: 655.
FEATURES              Location/Qualifiers
     source           1. .820
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:4288232"
                      /clone_lib="NIH_MGC_81"
                      /lab_host="DH10B (T1 phage-resistant)"
                      /note="Organ: muscle (skeletal); Vector: pDNR-LIB
                      (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI
                      (ggccattatggcc); 5' and 3' adaptors were used in cloning
                      as follows: 5' adaptor sequence:  5'-CACGGCCATTATGGCC-3'
                      and 3' adaptor sequence:
                      5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                      C, or G and N = A, C,  G, or T).   Average insert size
                      1.55 kb (range 1.0-4.0 kb).  15/15 colonies contained
                      inserts by PCR. This library was enriched for full-length
                      clones and was constructed by Clontech  Laboratories (Palo
                      Alto, CA)."
BASE COUNT      250 a    136 c    156 g    278 t
ORIGIN


  Query Match             1.4%;  Score 21;  DB 150;  Length 820;
  Best Local Similarity   100.0%;  Pred. No. 3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    818 taaatctattgaatggcattt 838
          |||||||||||||||||||||
Db    642 TAAATCTATTGAATGGCATTT 662


RESULT   8
BB582897/c
LOCUS       BB582897      250 bp    mRNA            EST       30-NOV-2000
DEFINITION  BB582897 RIKEN full-length enriched, adult male cecum Mus musculus
            cDNA clone 9130201G04 5', mRNA sequence.
ACCESSION   BB582897
VERSION     BB582897.1  GI:11479441
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 250)
  AUTHORS   Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T., Carninci,P.
            , Hanagaki,T., Hayatsu,N., Hiraoka,T., Hirozane,T., Hodoyama,Y.,
            Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J., Kojima,Y., Konno
            ,H., Kusakabe,M., Matsuyama,T., Miyazaki,A., Nakamura,M., Nishi,K.,
            Nomura,K., Numazaki,R., Okazaki,Y., Okido,T., Owa,C., Sakai,C.,
            Sakai,K., Sasaki,D., Sato,K., Shibata,K., Shibata,Y., Shinagawa,A.,
            Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Tanaka
            ,T., Toya,T., Watahiki,A., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshiki,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Aizawa,K. et al. 2000)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Genome Exploration Research Group, Life Science Tsukuba Center,
            Genome Science Laboratory
            The Institute of Physical and Chemical Research (RIKEN), Genomic
            Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: +81-298-36-9013
            Fax: +81-298-36-9098
            Email: genome-res@rtc.riken.go.jp,
            URL:http://genome.rtc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
            Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
            Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
            Carninci,P. and Hayashizaki,Y.
            High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
            Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES              Location/Qualifiers
     source           1. .250
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="9130201G04"
                      /clone_lib="RIKEN full-length enriched, adult male cecum"
                      /sex="male"
                      /tissue_type="cecum"
                      /dev_stage="adult"
```

/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"
BASE COUNT 65 a 65 c 72 g 48 t
ORIGIN

Query Match 1.3%; Score 20; DB 161; Length 250;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 239 gtgtttgcagcgtgcggccc 258
||||||||||||||||||||
Db 32 GTGTTTGCAGCGTGCGGCCC 13

RESULT 9
AA138114/c
LOCUS AA138114 353 bp mRNA EST 09-FEB-1997
DEFINITION mq75d06.r1 Stratagene mouse melanoma (#937312) Mus musculus cDNA clone IMAGE:584555 5', mRNA sequence.
ACCESSION AA138114
VERSION AA138114.1 GI:1700335
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 353)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:359203
Seq primer: -28m13 rev1 ET from Amersham.
FEATURES Location/Qualifiers
source 1. .353
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:584555"
/clone_lib="Stratagene mouse melanoma (#937312)"
/tissue_type="melanoma"
/dev_stage="M2 cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI ; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. From M2 cells, a highly metastatic derivative of the K-1735 (mouse) melanoma. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT 92 a 76 c 87 g 98 t
ORIGIN

Query Match 1.3%; Score 20; DB 2; Length 353;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1415 ggctccacattttctgggaa 1434
||||||||||||||||||||
Db 96 GGCTCCACATTTTCTGGGAA 77

RESULT 10
AL588417/c
LOCUS AL588417 380 bp mRNA EST 02-MAR-2001
DEFINITION AL588417 BP Chicken Brain Library Gallus gallus cDNA clone ROS072B02, mRNA sequence.
ACCESSION AL588417
VERSION AL588417.1 GI:13193451
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 380)
AUTHORS Murray,F.
TITLE BP Chicken Brain Library
JOURNAL Unpublished (2001)
COMMENT Contact: Frazer Murray
Dept. Genomics and Bioinformatics
Roslin Institute
Roslin, Midlothian, EH25 9PS, UK
Tel: +44 (0)131 527 4200
Fax: +44 (0)131 440 0434
Email: frazer.murray@bbsrc.ac.uk
GCGGCCGCTTTTTTTTTTTTTTTTT 3' Poly A RNA purchased from Clonetech (*6854-
Seq primer: T7.
FEATURES Location/Qualifiers
source 1. .380
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="ROS072B02"
/clone_lib="BP Chicken Brain Library"
/tissue_type="Brain"
/dev_stage="Unknown"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. 5' adaptor sequence: 5' TCGACCTCGAG 3' ; 3' adaptor sequence: 5' GCGGCCGCTTTTTTTTTTTTTTTTT 3' Poly A RNA purchased from Clonetech (*6854-1)"
BASE COUNT 75 a 114 c 139 g 51 t 1 others
ORIGIN

Query Match 1.3%; Score 20; DB 106; Length 380;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 834 catttgctctccagctggtt 853
||||||||||||||||||||
Db 289 CATTTGCTCTCCAGCTGGTT 270

RESULT 11

```
AQ667727
LOCUS       AQ667727      497 bp    DNA             GSS       24-JUN-1999
DEFINITION  HS_2137_B2_D10_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2137 Col=20 Row=H, DNA sequence.
ACCESSION   AQ667727
VERSION     AQ667727.1  GI:5200473
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 497)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2137  row: H  column: 20
            Seq primer: M13 Reverse
            Class: BAC ends
            High quality sequence stop: 497.
FEATURES             Location/Qualifiers
     source          1. .497
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=2137 Col=20 Row=H"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      146 a     91 c    118 g    140 t      2 others
ORIGIN


  Query Match             1.3%;  Score 20;  DB 231;  Length 497;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1423 attttctgggaatgggtcgg 1442
           ||||||||||||||||||||
Db     42 ATTTTCTGGGAATGGGTCGG 61


RESULT  12
BF318468/c
LOCUS       BF318468      518 bp    mRNA            EST       21-NOV-2000
DEFINITION  ux10c05.x1 Soares_thymus_2NbMT Mus musculus cDNA clone
            IMAGE:3471080 3', mRNA sequence.
ACCESSION   BF318468
VERSION     BF318468.1  GI:11266979
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 518)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: ux10c05.y1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1390440
            High quality sequence stop: 476.
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3471080"
                     /clone_lib="Soares_thymus_2NbMT"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
                     rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      134 a    129 c    101 g    154 t
ORIGIN


  Query Match             1.3%;  Score 20;  DB 147;  Length 518;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1340 gttacctacagttcctgcgg 1359
           ||||||||||||||||||||
Db    496 GTTACCTACAGTTCCTGCGG 477


RESULT  13
AZ739776/c
LOCUS       AZ739776      536 bp    DNA             GSS       25-JAN-2001
DEFINITION  RPCI-24-157E6.TV RPCI-24 Mus musculus genomic clone RPCI-24-157E6,
            DNA sequence.
ACCESSION   AZ739776
VERSION     AZ739776.1  GI:12510197
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 536)
  AUTHORS   Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-24-157E6.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 157  row: E  column: 6
            Seq primer: T7
            Class: BAC ends.
```

```
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-24-157E6"
                     /clone_lib="RPCI-24"
                     /sex="Male"
                     /cell_type="Spleen/Brain"
                     /note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1;
                     RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
                     library was  cloned in the pTARBAC1 cloning vector at the
                     BamH1 sites using MboI  partially digested male C57BL/6J
                     DNA."
BASE COUNT      194 a     85 c     91 g    166 t
ORIGIN


  Query Match             1.3%;  Score 20;  DB 248;  Length 536;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1179 cctgacagccatttacatta 1198
           ||||||||||||||||||||
Db    108 CCTGACAGCCATTTACATTA 89


RESULT  14
BG081170/c
LOCUS       BG081170      585 bp    mRNA            EST       26-JAN-2001
DEFINITION  H3061H07-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
            H3061H07 5', mRNA sequence.
ACCESSION   BG081170
VERSION     BG081170.1  GI:12563738
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
            ,T.S., Carter,M.G. and Ko,M.S.H.
  TITLE     Verification and initial annotation of NIA mouse 15K cDNA clone set
  JOURNAL   Unpublished (2001)
COMMENT     Other_ESTs: H3061H07-3
            Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
            Plate: H3061  row: H  column: 07
            Seq primer: -21M13 Reverse
            High quality sequence stop: 585
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="H3061H07"
                     /clone_lib="NIA Mouse 15K cDNA Clone Set"
                     /sex="Clones arrayed from a variety of cDNA libraries"
                     /dev_stage="Clones arrayed from a variety of cDNA
                     libraries"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     clone is among a rearrayed set of 15,247 clones from 11
                     embryo  cDNA libraries (including preimplantation stage
                     embryos from  unfertilized egg to blastocyst, embryonic
                     part of E7.5 embryos,  extraembryonic part of E7.5 embryos
                     , and E12.5 female mesonephros/gonad)  and one newborn
                     ovary cDNA library. Average insert size 1.5 kb. All
                     source libraries are cloned unidirectionally with Oligo(dT
                     )-Not primers.  References include: (1) Genome-wide
                     expression profiling of mid-gestation  placenta and embryo
                     using a 15,000 mouse developmental cDNA microarray,  2000,
                     Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
                     Large-scale cDNA  analysis reveals phased gene expression
                     patterns during preimplantation  mouse develolpment, 2000,
                     Development, 127: 1737-1749; (3) Genome-wide  mapping of
                     unselected transcripts from extraembryonic tissue of
                     7.5-day  mouse embryos reveals enrichment in the t-complex
                     and under-representation  on the X chromosome, 1998, Hum
                     Mol Genet 7: 1967-1978."
BASE COUNT      169 a    137 c    129 g    150 t
ORIGIN


  Query Match             1.3%;  Score 20;  DB 173;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1415 ggctccacattttctgggaa 1434
           ||||||||||||||||||||
Db     27 GGCTCCACATTTTCTGGGAA 8


RESULT  15
AZ739770/c
LOCUS       AZ739770      591 bp    DNA             GSS       25-JAN-2001
DEFINITION  RPCI-24-157E6.TJ RPCI-24 Mus musculus genomic clone RPCI-24-157E6,
            DNA sequence.
ACCESSION   AZ739770
VERSION     AZ739770.1  GI:12510183
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-24-157E6.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 157  row: E  column: 6
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-24-157E6"
                     /clone_lib="RPCI-24"
                     /sex="Male"
                     /cell_type="Spleen/Brain"
                     /note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1;
                     RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
                     library was  cloned in the pTARBAC1 cloning vector at the
                     BamH1 sites using MboI  partially digested male C57BL/6J
```

DNA."
BASE COUNT 223 a 95 c 100 g 173 t
ORIGIN

Query Match 1.3%; Score 20; DB 248; Length 591;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1179 cctgacagccatttacatta 1198
          ||||||||||||||||||||
Db    258 CCTGACAGCCATTTACATTA 239
```

Search completed: October 30, 2001, 12:28:17
Job time: 5674 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 15:21:54 ; Search time 19.36 Seconds
(without alignments)
1333.978 Million cell updates/sec

Title: US-09-619-049-855
Perfect score: 426
Sequence: 1 MRLPWVIFCTVLLLIFTNNS.....NKFTNEEEGRINCINFDSQT 426

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A_Geneseq_0601:*
1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	8	1.9	144	21	AAG28617	Arabidopsis thalia
2	8	1.9	284	19	AAW72112	HSV-2 strain SB5 C
3	8	1.9	785	19	AAW72195	HSV-2 strain SB5 C
4	8	1.9	3636	21	AAG47201	Arabidopsis thalia
5	8	1.9	3712	21	AAG47200	Arabidopsis thalia
6	8	1.9	3738	21	AAG47199	Arabidopsis thalia
7	7	1.6	13	22	AAB72418	Peptide used in th
8	7	1.6	97	21	AAB38359	Human secreted pro
9	7	1.6	138	22	AAB80170	Corynebacterium gl
10	7	1.6	143	21	AAB44011	Human cancer assoc
11	7	1.6	145	21	AAY58586	Sorangium cellulos
12	7	1.6	179	21	AAG19392	Arabidopsis thalia
13	7	1.6	179	21	AAG32444	Arabidopsis thalia
14	7	1.6	179	21	AAG47917	Arabidopsis thalia
15	7	1.6	180	21	AAG19391	Arabidopsis thalia
16	7	1.6	180	21	AAG32443	Arabidopsis thalia
17	7	1.6	180	21	AAG47916	Arabidopsis thalia
18	7	1.6	202	21	AAG26280	Arabidopsis thalia
19	7	1.6	203	21	AAG19390	Arabidopsis thalia
20	7	1.6	252	21	AAG21010	Arabidopsis thalia
21	7	1.6	252	21	AAG54018	Arabidopsis thalia
22	7	1.6	253	21	AAG21009	Arabidopsis thalia
23	7	1.6	253	21	AAG54017	Arabidopsis thalia
24	7	1.6	256	22	AAB71659	Human colon associ
25	7	1.6	260	20	AAY23987	TON1a protein enco
26	7	1.6	260	21	AAG21008	Arabidopsis thalia
27	7	1.6	260	21	AAG54016	Arabidopsis thalia
28	7	1.6	275	21	AAG26279	Arabidopsis thalia
29	7	1.6	311	21	AAG32442	Arabidopsis thalia
30	7	1.6	311	21	AAG47915	Arabidopsis thalia
31	7	1.6	334	22	AAB80169	Corynebacterium gl
32	7	1.6	354	20	AAY34932	Chlamydia pneumoni
33	7	1.6	359	19	AAW69598	Mouse G-protein co
34	7	1.6	360	19	AAW69597	Mouse G-protein co
35	7	1.6	372	21	AAG26278	Arabidopsis thalia
36	7	1.6	383	21	AAG20818	Arabidopsis thalia
37	7	1.6	385	20	AAY15082	Human protease-act
38	7	1.6	385	20	AAY50135	Human protease-act
39	7	1.6	385	20	AAY50138	Mutant human prote
40	7	1.6	385	20	AAY50139	Mutant human prote
41	7	1.6	385	21	AAY45036	Human protease act
42	7	1.6	396	20	AAY15081	Mouse protease-act
43	7	1.6	408	20	AAY50137	FLAG epitope-tagge
44	7	1.6	412	21	AAY95049	Candida albicans p
45	7	1.6	509	21	AAB13558	Streptomyces globi
46	7	1.6	709	22	AAB76859	Human lung tumour
47	7	1.6	1275	20	AAY29084	Human CIF130 prote
48	7	1.6	1275	21	AAY44262	Human cell cycle r
49	7	1.6	2010	21	AAB18218	Plasmodium falcipa
50	7	1.6	2466	20	AAY05844	Banana ripening fr
51	7	1.6	2502	21	AAB00117	M. tuberculosis po
52	6	1.4	15	20	AAW89508	HIV-1 protease inh
53	6	1.4	20	20	AAY29378	Final 3D profile w
54	6	1.4	24	19	AAW39950	Peptide effecting
55	6	1.4	48	21	AAB16567	Bacteriophage 44AH
56	6	1.4	51	21	AAG32517	Arabidopsis thalia
57	6	1.4	51	21	AAG52056	Arabidopsis thalia
58	6	1.4	53	20	AAY41408	Fragment of human
59	6	1.4	58	20	AAY17258	HLH domain of Myf-
60	6	1.4	66	20	AAY12512	Human 5' EST secre
61	6	1.4	67	21	AAG60324	Arabidopsis thalia
62	6	1.4	71	21	AAG32516	Arabidopsis thalia
63	6	1.4	71	21	AAG52055	Arabidopsis thalia
64	6	1.4	72	21	AAB44164	Human cancer assoc
65	6	1.4	73	20	AAY11794	Human 5' EST secre
66	6	1.4	75	21	AAG03040	Human secreted pro
67	6	1.4	76	20	AAW92957	B. subtilis SecG p
68	6	1.4	80	21	AAB06190	Caenorhabditis ele
69	6	1.4	86	21	AAG18974	Zea mays protein f
70	6	1.4	87	21	AAG08814	Arabidopsis thalia
71	6	1.4	87	21	AAG32515	Arabidopsis thalia
72	6	1.4	87	21	AAG52054	Arabidopsis thalia
73	6	1.4	93	21	AAG35344	Zea mays protein f
74	6	1.4	94	21	AAG15835	Arabidopsis thalia
75	6	1.4	98	21	AAG09345	Arabidopsis thalia
76	6	1.4	102	21	AAG56132	Arabidopsis thalia
77	6	1.4	102	21	AAG60366	Arabidopsis thalia
78	6	1.4	104	21	AAB13217	Caenorhabditis ele
79	6	1.4	109	21	AAG18973	Zea mays protein f
80	6	1.4	109	21	AAG59571	Arabidopsis thalia
81	6	1.4	112	22	AAB87628	Bovine mammary tis
82	6	1.4	114	21	AAG08813	Arabidopsis thalia
83	6	1.4	115	20	AAY32177	Chlamydia trachoma
84	6	1.4	116	20	AAY34893	Chlamydia pneumoni

85	6	1.4	116	21	AAG35343	Zea mays protein f
86	6	1.4	116	21	AAG40902	Zea mays protein f
87	6	1.4	117	21	AAG41072	Zea mays protein f
88	6	1.4	126	21	AAG59570	Arabidopsis thalia
89	6	1.4	128	21	AAG01437	Human secreted pro
90	6	1.4	129	21	AAG59569	Arabidopsis thalia
91	6	1.4	130	20	AAY36965	Chlamydia trachoma
92	6	1.4	130	21	AAB57111	Human prostate can
93	6	1.4	130	21	AAG01415	Human secreted pro
94	6	1.4	131	20	AAY60550	Human normal bladd
95	6	1.4	134	20	AAY25604	Lolium sp. allerge
96	6	1.4	134	21	AAG18928	Zea mays protein f
97	6	1.4	137	20	AAY41405	Fragment of human
98	6	1.4	138	21	AAG56131	Arabidopsis thalia
99	6	1.4	138	21	AAG60365	Arabidopsis thalia
100	6	1.4	139	21	AAG40901	Zea mays protein f
101	6	1.4	140	21	AAG08812	Arabidopsis thalia
102	6	1.4	141	21	AAB40828	Human ORFX ORF592
103	6	1.4	145	21	AAG52031	Arabidopsis thalia
104	6	1.4	146	14	AAR39717	Human fetal gamma
105	6	1.4	146	19	AAW73391	Human haemoglobin
106	6	1.4	146	19	AAW73393	Human haemoglobin
107	6	1.4	146	21	AAG56130	Arabidopsis thalia
108	6	1.4	146	21	AAG60364	Arabidopsis thalia
109	6	1.4	146	21	AAY50796	Human calmodulin/N
110	6	1.4	146	22	AAB66771	Human hemoglobin f
111	6	1.4	147	21	AAG03988	Human secreted pro
112	6	1.4	152	21	AAG35342	Zea mays protein f
113	6	1.4	154	21	AAY74339	Neisseria gonorrhe
114	6	1.4	154	21	AAY74340	Neisseria meningit
115	6	1.4	154	21	AAY74341	Neisseria meningit
116	6	1.4	155	19	AAW75101	Human secreted pro
117	6	1.4	156	20	AAY14050	G. oxydans D-sorbi
118	6	1.4	161	21	AAG52030	Arabidopsis thalia
119	6	1.4	170	20	AAY35732	Chlamydia pneumoni
120	6	1.4	174	21	AAG60079	Arabidopsis thalia
121	6	1.4	175	14	AAR33390	Albumin-type prote
122	6	1.4	176	21	AAG60078	Arabidopsis thalia
123	6	1.4	179	21	AAG60077	Arabidopsis thalia
124	6	1.4	179	21	AAY32469	DNA encoding felin
125	6	1.4	182	18	AAW28250	Amino acid sequenc
126	6	1.4	189	21	AAB32022	Human secreted pro
127	6	1.4	191	18	AAW20418	H. pylori surface
128	6	1.4	193	15	AAR50121	CD27L. Homo sapie
129	6	1.4	193	21	AAG18033	Arabidopsis thalia
130	6	1.4	197	21	AAG52029	Arabidopsis thalia
131	6	1.4	207	20	AAY35596	Chlamydia pneumoni
132	6	1.4	207	20	AAW94667	Streptococcus pneu
133	6	1.4	213	17	AAW06303	Borrelia lonestari
134	6	1.4	213	20	AAY16592	Human glutathione
135	6	1.4	213	20	AAW92526	Human GSTH protein
136	6	1.4	214	20	AAY37004	Chlamydia trachoma
137	6	1.4	216	15	AAR53971	CD27-L type II tra
138	6	1.4	216	15	AAR50122	sCD27L-3. Homo sa
139	6	1.4	216	19	AAW41180	CD27 ligand. Homo
140	6	1.4	217	21	AAB01361	1-deoxy-D-xylulose
141	6	1.4	219	21	AAG49000	Arabidopsis thalia
142	6	1.4	223	21	AAG48999	Arabidopsis thalia
143	6	1.4	225	21	AAG25205	Arabidopsis thalia
144	6	1.4	227	22	AAB63883	Human prostate can
145	6	1.4	228	21	AAB42047	Human ORFX ORF1811
146	6	1.4	228	21	AAG49931	Arabidopsis thalia
147	6	1.4	229	20	AAW94666	Streptococcus pneu
148	6	1.4	230	21	AAB56877	Human prostate can
149	6	1.4	232	21	AAG49930	Arabidopsis thalia
150	6	1.4	241	20	AAY29557	Human lung tumour
151	6	1.4	241	21	AAB44480	Human lung tumour-
152	6	1.4	245	21	AAY94627	HIF-1alpha variant
153	6	1.4	246	21	AAG25204	Arabidopsis thalia
154	6	1.4	247	18	AAY11115	S. pneumoniae prob
155	6	1.4	256	21	AAB01291	Mature cobra venom
156	6	1.4	256	21	AAB01293	Altered cobra veno
157	6	1.4	256	22	AAB71828	Infectious salmon
158	6	1.4	261	21	AAB25372	Pinus radiata cell
159	6	1.4	262	21	AAG12782	Arabidopsis thalia
160	6	1.4	262	21	AAG53105	Arabidopsis thalia
161	6	1.4	263	19	AAW56277	Babesia microti BM
162	6	1.4	263	20	AAY24339	Babesia microti an
163	6	1.4	263	21	AAB30188	B. microti BMNI-1
164	6	1.4	264	20	AAY14820	PanB amino acid se
165	6	1.4	264	20	AAW87973	A lectin derived p
166	6	1.4	267	19	AAW56293	Babesia microti BM
167	6	1.4	267	20	AAY24349	Babesia microti an
168	6	1.4	267	21	AAB30198	B. microti BMNI-13
169	6	1.4	267	21	AAY70223	Human RNA-associat
170	6	1.4	268	21	AAG22455	Arabidopsis thalia
171	6	1.4	269	21	AAY88289	Murine TANGO 188 p
172	6	1.4	271	18	AAW19612	Human growth hormo
173	6	1.4	271	18	AAW19219	Human growth hormo
174	6	1.4	271	21	AAG22454	Arabidopsis thalia
175	6	1.4	272	18	AAW20996	H. pylori surface
176	6	1.4	275	14	AAR43390	Sequence specific
177	6	1.4	275	17	AAR98993	Neisseria gonorrho
178	6	1.4	275	21	AAY67400	N. gonorrheae spec
179	6	1.4	276	21	AAG48018	Arabidopsis thalia
180	6	1.4	282	21	AAG48160	Arabidopsis thalia
181	6	1.4	284	21	AAB42348	Human ORFX ORF2112
182	6	1.4	285	21	AAG12781	Arabidopsis thalia
183	6	1.4	285	21	AAG48159	Arabidopsis thalia
184	6	1.4	285	21	AAG53104	Arabidopsis thalia
185	6	1.4	288	21	AAG31787	Arabidopsis thalia
186	6	1.4	289	21	AAG17475	Arabidopsis thalia
187	6	1.4	289	21	AAG31786	Arabidopsis thalia
188	6	1.4	290	21	AAG48998	Arabidopsis thalia
189	6	1.4	290	22	AAB31632	Amino acid sequenc
190	6	1.4	291	21	AAB28178	Protein expressed
191	6	1.4	291	21	AAB28179	Protein expressed
192	6	1.4	291	22	AAB69092	Mouse GN protein s
193	6	1.4	291	22	AAB31156	Amino acid sequenc
194	6	1.4	294	19	AAW56286	Babesia microti BM
195	6	1.4	294	20	AAY24363	Babesia microti an
196	6	1.4	294	21	AAB30212	B. microti BMNI-7
197	6	1.4	294	21	AAG12780	Arabidopsis thalia
198	6	1.4	294	21	AAG13958	Arabidopsis thalia
199	6	1.4	294	21	AAG46341	Arabidopsis thalia
200	6	1.4	294	21	AAG53103	Arabidopsis thalia
201	6	1.4	295	15	AAR48758	Human cytomegalovi
202	6	1.4	295	17	AAW02730	Human cytomegalovi
203	6	1.4	299	21	AAG49929	Arabidopsis thalia
204	6	1.4	299	22	AAB87731	Human T2R01 amino
205	6	1.4	301	21	AAY77961	A. thaliana enviro
206	6	1.4	303	19	AAW56285	Babesia microti BM
207	6	1.4	303	20	AAY24344	Babesia microti an
208	6	1.4	303	21	AAB30193	B. microti BMNI-6
209	6	1.4	307	22	AAB79897	Corynebacterium gl
210	6	1.4	309	21	AAY57447	Mouse homologue of
211	6	1.4	310	19	AAW56278	Babesia microti BM
212	6	1.4	310	20	AAY24340	Babesia microti an
213	6	1.4	310	21	AAB30189	B. microti BMNI-2
214	6	1.4	311	21	AAG48017	Arabidopsis thalia
215	6	1.4	313	19	AAW75159	Human secreted pro
216	6	1.4	318	14	AAR41527	Human A3 adenosine
217	6	1.4	318	16	AAR93992	Human striatum A3
218	6	1.4	318	16	AAR87658	Human adenosine re
219	6	1.4	318	16	AAR84195	Human A3 adenosine
220	6	1.4	318	20	AAW82313	Human A3 adenosine
221	6	1.4	320	19	AAW80668	S. pneumoniae nucl
222	6	1.4	320	21	AAY49370	S. pneumoniae nrdF
223	6	1.4	321	22	AAB65977	Human secreted pro
224	6	1.4	321	22	AAB65978	Human secreted pro
225	6	1.4	321	22	AAB64717	Gene 5 human secre
226	6	1.4	321	22	AAB64718	Human secreted pro
227	6	1.4	322	17	AAW03688	Leishmania chagasi
228	6	1.4	324	10	AAP90552	Hamster beta-2 -ad
229	6	1.4	324	20	AAW80360	The small subunit
230	6	1.4	324	21	AAB42812	Human ORFX ORF2576

231	6	1.4	324	21	AAG31785	Arabidopsis thalia
232	6	1.4	324	21	AAG56724	Arabidopsis thalia
233	6	1.4	324	21	AAG60166	Arabidopsis thalia
234	6	1.4	326	18	AAW21022	H. pylori cytoplas
235	6	1.4	330	21	AAY94628	HIF-lalpha variant
236	6	1.4	330	21	AAY75966	Rat skin cell prot
237	6	1.4	330	21	AAY76023	Rat skin cell prot
238	6	1.4	330	22	AAB55905	Skin cell protein,
239	6	1.4	330	22	AAB55962	Skin cell protein,
240	6	1.4	331	21	AAG17474	Arabidopsis thalia
241	6	1.4	337	20	AAY37741	Protein which is s
242	6	1.4	339	21	AAG17473	Arabidopsis thalia
243	6	1.4	339	21	AAG51245	Arabidopsis thalia
244	6	1.4	347	21	AAB43540	Human cancer assoc
245	6	1.4	349	19	AAW44368	Aspergillus nidula
246	6	1.4	349	21	AAY69293	A canine growth ho
247	6	1.4	354	21	AAG41450	Arabidopsis thalia
248	6	1.4	356	21	AAG22453	Arabidopsis thalia
249	6	1.4	361	18	AAW19217	Human growth hormo
250	6	1.4	362	18	AAW19610	Human growth hormo
251	6	1.4	364	18	AAW19613	Rat growth hormone
252	6	1.4	364	18	AAW19220	Rat growth hormone
253	6	1.4	364	19	AAW65453	Drug binding prote
254	6	1.4	364	19	AAW40061	Human MAP kinase p
255	6	1.4	364	20	AAY42424	p38-beta mutant.
256	6	1.4	364	21	AAY54565	A mouse growth hor
257	6	1.4	364	22	AAB79284	Corynebacterium gl
258	6	1.4	366	21	AAY90632	Human G protein-co
259	6	1.4	366	21	AAY90666	Human mutant G pro
260	6	1.4	366	21	AAY70345	Human G protein-co
261	6	1.4	367	19	AAW56301	Babesia microti an
262	6	1.4	367	19	AAW56279	Babesia microti BM
263	6	1.4	367	20	AAY24341	Babesia microti an
264	6	1.4	367	20	AAY24368	Babesia microti an
265	6	1.4	367	21	AAB30190	B. microti BMNI-3
266	6	1.4	367	21	AAB30215	B. microti BMNI-3
267	6	1.4	373	18	AAW06559	HIF-1 alpha delta-
268	6	1.4	378	21	AAB53159	Macaca mulatta rha
269	6	1.4	381	21	AAG05735	Arabidopsis thalia
270	6	1.4	382	21	AAG48016	Arabidopsis thalia
271	6	1.4	384	22	AAB65609	Novel protein kina
272	6	1.4	385	21	AAG13957	Arabidopsis thalia
273	6	1.4	385	21	AAG46340	Arabidopsis thalia
274	6	1.4	391	22	AAB60850	Mevalonate pathway
275	6	1.4	392	21	AAB26438	Drosophila melanog
276	6	1.4	393	21	AAB20941	Drosophila odorant
277	6	1.4	396	21	AAB15501	Bacillus subtilis
278	6	1.4	396	21	AAB15502	Bacillus subtilis
279	6	1.4	396	21	AAB15503	Bacillus subtilis
280	6	1.4	399	21	AAY97197	A. thaliana partia
281	6	1.4	401	21	AAG26448	Arabidopsis thalia
282	6	1.4	402	21	AAG56723	Arabidopsis thalia
283	6	1.4	402	21	AAG60165	Arabidopsis thalia
284	6	1.4	405	21	AAG26447	Arabidopsis thalia
285	6	1.4	406	21	AAB11412	A. thaliana DXPRI
286	6	1.4	406	21	AAB01370	1-deoxy-D-xylulose
287	6	1.4	407	21	AAG41449	Arabidopsis thalia
288	6	1.4	410	21	AAG26446	Arabidopsis thalia
289	6	1.4	411	19	AAW80133	Arthrobacter sp. c
290	6	1.4	414	20	AAY30209	Aa human cytosolic
291	6	1.4	415	21	AAG56722	Arabidopsis thalia
292	6	1.4	415	21	AAG60164	Arabidopsis thalia
293	6	1.4	419	21	AAG41448	Arabidopsis thalia
294	6	1.4	420	21	AAB11413	A. thaliana DXPRI
295	6	1.4	423	20	AAY76514	Human ovarian tumo
296	6	1.4	423	21	AAG48158	Arabidopsis thalia
297	6	1.4	426	20	AAY59993	Human endometrium
298	6	1.4	426	21	AAB56924	Human prostate can
299	6	1.4	429	21	AAG13956	Arabidopsis thalia
300	6	1.4	429	21	AAG46339	Arabidopsis thalia
301	6	1.4	431	20	AAY34509	Porphorymonas ging
302	6	1.4	432	19	AAW54233	Human cytokeratin
303	6	1.4	439	19	AAW74484	Peniophora phytase
304	6	1.4	439	19	AAW62858	Peniophora lycii 6
305	6	1.4	439	20	AAY39895	P. lycii phytase p
306	6	1.4	453	21	AAY68799	Amino acid sequenc
307	6	1.4	455	11	AAR06257	Chicken acetylchol
308	6	1.4	457	12	AAR15471	BBRP42. Bacillus
309	6	1.4	458	21	AAY97013	S. cerevisiae esse
310	6	1.4	459	21	AAB58761	Breast and ovarian
311	6	1.4	460	15	AAR49667	Human nucleobindin
312	6	1.4	460	20	AAY34382	Porphorymonas ging
313	6	1.4	460	22	AAB68506	Human GTP-binding
314	6	1.4	464	17	AAR92214	SAK-b serine-threo
315	6	1.4	466	21	AAY83044	F-box protein FBP-
316	6	1.4	468	20	AAY38583	Neisseria gonorrho
317	6	1.4	470	22	AAB20327	Human protein phos
318	6	1.4	471	17	AAW03756	Maize UDP-glucose
319	6	1.4	471	21	AAG23141	Arabidopsis thalia
320	6	1.4	471	21	AAG46521	Arabidopsis thalia
321	6	1.4	471	21	AAG50541	Arabidopsis thalia
322	6	1.4	472	21	AAB01367	1-deoxy-D-xylulose
323	6	1.4	472	21	AAG51244	Arabidopsis thalia
324	6	1.4	473	21	AAB01362	1-deoxy-D-xylulose
325	6	1.4	473	21	AAB01369	1-deoxy-D-xylulose
326	6	1.4	473	21	AAG23140	Arabidopsis thalia
327	6	1.4	473	21	AAG46520	Arabidopsis thalia
328	6	1.4	475	19	AAW44741	Human neuronal str
329	6	1.4	475	21	AAY97196	Mentha piperita 1-
330	6	1.4	475	21	AAB01371	1-deoxy-D-xylulose
331	6	1.4	476	21	AAB30296	Arabidopsis D-1-de
332	6	1.4	477	15	AAR66067	Phosphorylated p66
333	6	1.4	477	21	AAY95658	Arabidopsis herbic
334	6	1.4	477	21	AAY95659	Arabidopsis herbic
335	6	1.4	477	22	AAB31985	1-deoxy-D-xylulose
336	6	1.4	477	22	AAB31986	1-deoxy-D-xylulose
337	6	1.4	479	19	AAW57901	Protein of clone C
338	6	1.4	479	21	AAB08443	Amino acid sequenc
339	6	1.4	479	21	AAG50540	Arabidopsis thalia
340	6	1.4	486	19	AAW38455	Yeast RNA-binding
341	6	1.4	487	12	AAR14149	Human alpha 2 beta
342	6	1.4	487	18	AAW11804	Human alpha-2b adr
343	6	1.4	488	20	AAY36300	Human secreted pro
344	6	1.4	489	21	AAG40251	Arabidopsis thalia
345	6	1.4	490	21	AAG23139	Arabidopsis thalia
346	6	1.4	490	21	AAG31141	Arabidopsis thalia
347	6	1.4	490	21	AAG46519	Arabidopsis thalia
348	6	1.4	491	21	AAG40250	Arabidopsis thalia
349	6	1.4	491	21	AAY59193	Mouse alpha-2 anti
350	6	1.4	496	21	AAY70762	Spider protein dis
351	6	1.4	498	20	AAY38584	Neisseria gonorrho
352	6	1.4	498	21	AAG51243	Arabidopsis thalia
353	6	1.4	498	21	AAY74707	Neisseria gonorrhe
354	6	1.4	499	21	AAB01363	1-deoxy-D-xylulose
355	6	1.4	499	21	AAG29414	Arabidopsis thalia
356	6	1.4	501	17	AAR90578	Mouse TRAF2. Mus
357	6	1.4	502	20	AAY37519	Protein involved i
358	6	1.4	503	20	AAY15103	Modified Interfero
359	6	1.4	505	9	AAP80885	Sequence encoded b
360	6	1.4	505	19	AAW72992	HIV isolate LAV.MA
361	6	1.4	505	21	AAY69159	Peptide QQ3351-BF0
362	6	1.4	509	18	AAW18786	Neisseria adhesion
363	6	1.4	510	21	AAG40249	Arabidopsis thalia
364	6	1.4	511	21	AAG29413	Arabidopsis thalia
365	6	1.4	529	18	AAW19001	Feline herpes viru
366	6	1.4	529	19	AAW55054	Sunflower leaf ant
367	6	1.4	529	19	AAW55053	Sunflower leaf ant
368	6	1.4	529	21	AAG31140	Arabidopsis thalia
369	6	1.4	529	22	AAB61787	Sunflower carbohyd
370	6	1.4	529	22	AAB61788	Sunflower carbohyd
371	6	1.4	532	16	AAR82937	Mouse Kv1.7 voltag
372	6	1.4	533	21	AAB19001	A partial microtub
373	6	1.4	537	21	AAY95062	Candida albicans p
374	6	1.4	540	22	AAB49401	Murine organic ani
375	6	1.4	542	20	AAY18035	S. cerevisiae INV9
376	6	1.4	543	16	AAR76569	Beetle luciferase

377	6	1.4	544	20	AAY06862	Mutant luciferase
378	6	1.4	544	20	AAY06863	Mutant luciferase
379	6	1.4	544	21	AAG31954	Arabidopsis thalia
380	6	1.4	545	20	AAY06860	Mutant luciferase
381	6	1.4	546	20	AAY06858	Mutant luciferase
382	6	1.4	546	20	AAY06859	Mutant luciferase
383	6	1.4	546	20	AAY06857	Mutant luciferase
384	6	1.4	553	21	AAB24225	Human vesicle asso
385	6	1.4	562	21	AAB18313	Plasmodium falcipa
386	6	1.4	569	20	AAY37113	Chlamydia trachoma
387	6	1.4	580	17	AAR94583	C. pneumoniae poly
388	6	1.4	580	18	AAW01742	C. pneumoniae 73 k
389	6	1.4	580	18	AAW11707	C. pneumoniae 73K
390	6	1.4	580	20	AAY34914	Chlamydia pneumoni
391	6	1.4	586	21	AAG30917	Arabidopsis thalia
392	6	1.4	591	21	AAY58904	Rhinocladiella atr
393	6	1.4	591	21	AAY58905	Rhinocladiella atr
394	6	1.4	597	21	AAY58114	Haemorrhagic enter
395	6	1.4	598	21	AAY58902	Exophiala spinifer
396	6	1.4	598	21	AAY58903	Rhinocladiella atr
397	6	1.4	600	21	AAY58900	Exophiala spinifer
398	6	1.4	600	21	AAY58901	Exophiala spinifer
399	6	1.4	600	21	AAY58913	Exophiala spinifer
400	6	1.4	600	21	AAY68849	A full length amin
401	6	1.4	612	21	AAG50539	Arabidopsis thalia
402	6	1.4	613	21	AAY94630	HIF-1alpha variant
403	6	1.4	615	21	AAG30916	Arabidopsis thalia
404	6	1.4	622	22	AAB48832	Trametes hirsuta p
405	6	1.4	623	9	AAP81339	Expression prod. o
406	6	1.4	632	14	AAR41232	GAT-3 transporter.
407	6	1.4	632	21	AAB06177	Caenorhabditis ele
408	6	1.4	633	21	AAG31342	Arabidopsis thalia
409	6	1.4	634	13	AAR23069	GL-7ACA aycylase.
410	6	1.4	636	21	AAB06178	Caenorhabditis ele
411	6	1.4	642	21	AAY84168	A variant of human
412	6	1.4	647	11	AAR06207	Plasmid FokI restr
413	6	1.4	648	20	AAY52834	Rhodobacter sphaer
414	6	1.4	652	21	AAY94629	HIF-1alpha variant
415	6	1.4	669	21	AAY84167	A variant of human
416	6	1.4	671	21	AAY74342	Neisseria gonorrhe
417	6	1.4	671	21	AAY74343	Neisseria meningit
418	6	1.4	671	21	AAY74344	Neisseria meningit
419	6	1.4	678	8	AAP70380	Plasmid pRF1 rat 1
420	6	1.4	678	8	AAP70578	NADPH cytochrome P
421	6	1.4	684	21	AAG30915	Arabidopsis thalia
422	6	1.4	694	21	AAG31953	Arabidopsis thalia
423	6	1.4	697	21	AAY84166	A variant of human
424	6	1.4	701	21	AAY84173	A variant of human
425	6	1.4	703	21	AAG31139	Arabidopsis thalia
426	6	1.4	710	21	AAY84172	A variant of human
427	6	1.4	712	22	AAB79540	Corynebacterium gl
428	6	1.4	715	20	AAY27348	Group B Streptococ
429	6	1.4	724	21	AAY84171	A variant of human
430	6	1.4	739	16	AAR77501	Malate synthase.
431	6	1.4	739	22	AAB79539	Corynebacterium gl
432	6	1.4	743	21	AAG31952	Arabidopsis thalia
433	6	1.4	747	21	AAB19002	A human microtubul
434	6	1.4	747	21	AAB29444	Human foetal liver
435	6	1.4	749	21	AAY84170	A variant of human
436	6	1.4	756	21	AAY94635	HIF-1alpha variant
437	6	1.4	771	16	AAR80495	Japanese oyster tr
438	6	1.4	781	19	AAW57887	Soybean raffinose
439	6	1.4	781	20	AAY30143	Amino acid sequenc
440	6	1.4	781	22	AAB49400	Soybean raffinose
441	6	1.4	784	20	AAY06618	Thermomyces lanugi
442	6	1.4	788	21	AAY85274	C. albicans CaNL25
443	6	1.4	789	21	AAY84169	A variant of human
444	6	1.4	793	20	AAY27347	Group B Streptococ
445	6	1.4	805	18	AAW06558	Hypoxia inducible
446	6	1.4	813	21	AAY94636	HIF-1alpha variant
447	6	1.4	823	21	AAB12746	Streptococcus pneu
448	6	1.4	825	22	AAU00026	Streptococcus pyog
449	6	1.4	825	22	AAU00029	Streptococcus pyog
450	6	1.4	826	18	AAW06557	Human hypoxia indu
451	6	1.4	826	19	AAW80418	Amino acid sequenc
452	6	1.4	826	20	AAY06289	Human transcriptio
453	6	1.4	826	21	AAY94640	Human hypoxia-indu
454	6	1.4	826	21	AAY69407	A wild type human
455	6	1.4	826	22	AAB76854	Human lung tumour
456	6	1.4	827	20	AAY34499	Porphorymonas ging
457	6	1.4	828	19	AAW64473	Human secreted pro
458	6	1.4	828	20	AAY34372	Porphorymonas ging
459	6	1.4	829	21	AAB58894	Breast and ovarian
460	6	1.4	829	21	AAY58918	GST-amino polyol a
461	6	1.4	829	21	AAY68854	A fusion of GST/mu
462	6	1.4	834	20	AAY34869	Chlamydia pneumoni
463	6	1.4	836	20	AAY37012	Protein involved i
464	6	1.4	836	21	AAB28532	Rice RNA-directed
465	6	1.4	852	19	AAW79065	FIV-NCSU1 clone JS
466	6	1.4	855	15	AAR58585	FIV UK8 envelope p
467	6	1.4	855	18	AAW23711	Feline immunodefic
468	6	1.4	856	18	AAW08436	Feline immunodefic
469	6	1.4	856	18	AAW01823	FIV envelope polyp
470	6	1.4	856	18	AAW07846	FIV envelope prote
471	6	1.4	870	15	AAR51246	Composite sequence
472	6	1.4	895	17	AAR91615	Trypanosoma cruzi
473	6	1.4	906	14	AAR44300	Human GluR1B recep
474	6	1.4	906	14	AAR42354	Human glutamate re
475	6	1.4	906	16	AAR84912	Human glutamate re
476	6	1.4	906	16	AAR84913	Human glutamate re
477	6	1.4	907	12	AAR11989	Glutamate receptor
478	6	1.4	915	20	AAY22124	T. cruzi PTc40 pro
479	6	1.4	967	21	AAG38847	Arabidopsis thalia
480	6	1.4	974	19	AAW59357	Human retinal dege
481	6	1.4	974	21	AAY93246	An Escherichia col
482	6	1.4	984	22	AAB79887	Corynebacterium gl
483	6	1.4	984	22	AAB79930	Corynebacterium gl
484	6	1.4	998	21	AAB25539	Eucalyptus grandis
485	6	1.4	998	21	AAB25556	Pinus radiata cell
486	6	1.4	1013	19	AAW54377	IBDV VP2-VP4-VP3 p
487	6	1.4	1016	21	AAG38846	Arabidopsis thalia
488	6	1.4	1020	21	AAG38845	Arabidopsis thalia
489	6	1.4	1024	21	AAG42350	Arabidopsis thalia
490	6	1.4	1073	7	AAP60569	Sequence of the R1
491	6	1.4	1099	17	AAW05177	Lats large tumour
492	6	1.4	1099	21	AAY70393	Drosophila melanog
493	6	1.4	1105	21	AAG42349	Arabidopsis thalia
494	6	1.4	1122	21	AAY32218	Drosophila PER B p
495	6	1.4	1132	9	AAP81338	Expression prod. o
496	6	1.4	1140	16	AAR72386	XAP-1, part of the
497	6	1.4	1144	9	AAP81334	Expression prod. o
498	6	1.4	1150	9	AAP81337	Expression prod. o
499	6	1.4	1150	9	AAP81335	Expression prod. o
500	6	1.4	1162	9	AAP81336	Expression prod. o
501	6	1.4	1175	21	AAY32219	Drosophila PER C p
502	6	1.4	1189	21	AAG42348	Arabidopsis thalia
503	6	1.4	1218	21	AAY32217	Drosophila PER A p
504	6	1.4	1233	22	AAB20125	Human sodium chann
505	6	1.4	1243	20	AAY06598	Human sodium chann
506	6	1.4	1243	22	AAB20126	Human sodium chann
507	6	1.4	1250	21	AAG38593	Arabidopsis thalia
508	6	1.4	1291	21	AAG38592	Arabidopsis thalia
509	6	1.4	1349	19	AAW59359	Human retinal dege
510	6	1.4	1371	21	AAG38591	Arabidopsis thalia
511	6	1.4	1422	16	AAR82068	Hepatitis GB virus
512	6	1.4	1422	21	AAB09038	Hepatitis GB virus
513	6	1.4	1541	18	AAW33361	Rat canalicular mu
514	6	1.4	1545	18	AAW33362	Human canalicular
515	6	1.4	1545	19	AAW55966	Human canalicular
516	6	1.4	1642	15	AAR63223	Cobra CVF1. Naja
517	6	1.4	1642	20	AAY23729	Cobra venom factor
518	6	1.4	1791	22	AAB20121	Human sodium chann
519	6	1.4	1856	20	AAY21805	Reading frame 3 pr
520	6	1.4	1856	21	AAY83274	Polypeptide encode
521	6	1.4	1979	21	AAB18171	Plasmodium falcipa
522	6	1.4	2037	21	AAB23834	Shewanella sp. SCR

523	6	1.4	2183	14	AAR39592	L protein of atten
524	6	1.4	2183	19	AAW48706	Measles virus Mont
525	6	1.4	2183	19	AAW48707	Measles virus Rube
526	6	1.4	2183	19	AAW48708	Measles virus Mora
527	6	1.4	2183	19	AAW48709	Measles virus Zagr
528	6	1.4	2183	19	AAW48710	Measles virus AIK-
529	6	1.4	2183	19	AAW48703	Measles virus Edmo
530	6	1.4	2183	19	AAW48704	Measles virus 1977
531	6	1.4	2183	19	AAW48705	Measles virus 1983
532	6	1.4	2210	21	AAB10040	LCMV L protein. L
533	6	1.4	2471	20	AAW88481	Candida albicans h
534	6	1.4	2707	18	AAW27161	Mouse receptor ME2
535	6	1.4	2864	16	AAR82072	Hepatitis GB virus
536	6	1.4	2864	21	AAB09268	Hepatitis GB virus
537	6	1.4	2864	22	AAB59171	Protein encoded by
538	5	1.2	5	12	AAR11938	Pentapeptide paral
539	5	1.2	5	15	AAR51518	Mimotope peptide #
540	5	1.2	5	17	AAR98632	Peptide 12 from 88
541	5	1.2	5	20	AAY31080	Non-crosslinked pr
542	5	1.2	6	9	AAP82440	Peptide unit of pr
543	5	1.2	6	22	AAB47109	Epitope #1 used in
544	5	1.2	7	14	AAR32669	Peptide heptad. S
545	5	1.2	7	14	AAR32672	Peptide heptad. S
546	5	1.2	7	14	AAR32675	Peptide heptad. S
547	5	1.2	7	19	AAW68518	MAb MCPC603 Vl CDR
548	5	1.2	8	16	AAR73477	Human TSH receptor
549	5	1.2	8	16	AAR73478	Human TSH receptor
550	5	1.2	8	16	AAR73312	Human TSH receptor
551	5	1.2	8	16	AAR73313	Human TSH receptor
552	5	1.2	8	16	AAR69836	Peptide comprising
553	5	1.2	8	21	AAY77688	TSG-6 binding pept
554	5	1.2	9	14	AAR43247	Cytotoxic T lympho
555	5	1.2	9	16	AAR87312	Plasmodium falcipa
556	5	1.2	9	16	AAR84769	MART-1 melanoma an
557	5	1.2	9	20	AAY47588	Immunogenic peptid
558	5	1.2	9	20	AAY10563	HLA Class I motif
559	5	1.2	9	21	AAY95894	Human MHC class I
560	5	1.2	9	22	AAB68699	Major outer membra
561	5	1.2	9	22	AAB46246	HPV type 16 cytoto
562	5	1.2	10	14	AAR33482	Cytotoxic T-lympho
563	5	1.2	10	15	AAR57928	Randomly generated
564	5	1.2	10	16	AAR84779	MART-1 melanoma an
565	5	1.2	10	20	AAY40649	S2 derivative #11,
566	5	1.2	10	20	AAY47590	Immunogenic peptid
567	5	1.2	10	21	AAB29988	Scaffold protein S
568	5	1.2	10	21	AAB39134	Human secreted pro
569	5	1.2	10	22	AAB66835	Metal ion binding
570	5	1.2	11	17	AAW02108	Guinea pig L-aspar
571	5	1.2	11	18	AAW23272	Tumour necrosis fa
572	5	1.2	11	20	AAW95470	LS170 polypeptide
573	5	1.2	11	21	AAB30281	CD4+ T-cell activa
574	5	1.2	11	21	AAB30289	CD4+ T-cell activa
575	5	1.2	12	12	AAR13012	Leuko T-1. Synthe
576	5	1.2	12	18	AAW36370	Antimicrobial prot
577	5	1.2	12	18	AAW36357	Antimicrobial prot
578	5	1.2	12	20	AAY07320	Collagen assembly
579	5	1.2	12	21	AAY88104	Oregon green 514 b
580	5	1.2	12	21	AAY88154	Oregon green 514 b
581	5	1.2	12	21	AAY87933	Peptide #7 from pr
582	5	1.2	12	21	AAY87935	Peptide #9 from pr
583	5	1.2	12	21	AAY54435	Peptide which mimi
584	5	1.2	12	22	AAB36509	Drosophila Hsp83 r
585	5	1.2	13	16	AAR62959	Chrysanthemyl-diph
586	5	1.2	13	18	AAW36375	Antimicrobial prot
587	5	1.2	13	18	AAW35618	Antimicrobial pept
588	5	1.2	13	19	AAW70384	human regeneration
589	5	1.2	13	19	AAW51832	Rana temporaria te
590	5	1.2	13	20	AAY39599	CTLA-4 VLD CDR3 re
591	5	1.2	13	21	AAB30271	CD4+ T-cell activa
592	5	1.2	13	21	AAB18741	Amino acid sequenc
593	5	1.2	13	21	AAB34884	Human secreted pro
594	5	1.2	13	21	AAB16462	Cylic peptide that
595	5	1.2	14	15	AAR53372	N-terminal of 82 k
596	5	1.2	14	18	AAW31849	Human release fact
597	5	1.2	14	19	AAW41594	Rat FRAG1 protein
598	5	1.2	14	22	AAB88180	CD66 peptide CD66f
599	5	1.2	15	15	AAR53373	N-terminal of muri
600	5	1.2	15	15	AAR47675	HIV epitope #8. H
601	5	1.2	15	16	AAR64821	HPV-16 L-protein(3
602	5	1.2	15	16	AAR64822	HPV-16 L-protein(3
603	5	1.2	15	17	AAW49183	Human leucocyte an
604	5	1.2	15	17	AAW49162	Human leucocyte an
605	5	1.2	15	18	AAW38933	Peptide resembling
606	5	1.2	15	19	AAW44130	Chlamydia mitochon
607	5	1.2	15	20	AAY08382	Cysteine noose lib
608	5	1.2	15	20	AAW73715	M. tuberculosis an
609	5	1.2	15	20	AAW73717	M. tuberculosis an
610	5	1.2	15	20	AAW73827	M. tuberculosis an
611	5	1.2	15	20	AAW73825	M. tuberculosis an
612	5	1.2	15	21	AAB19575	Fibrinogen-derived
613	5	1.2	15	21	AAY84182	Amino acid sequenc
614	5	1.2	15	22	AAB89984	HIV gp120 protein
615	5	1.2	15	22	AAB90001	HIV gp120 protein
616	5	1.2	16	12	AAR14158	CS3-related peptid
617	5	1.2	16	15	AAR52187	Human light chain
618	5	1.2	16	15	AAR54015	P-glycoprotein pep
619	5	1.2	16	18	AAW31062	Mugwort pollen all
620	5	1.2	16	21	AAB37048	Bcl2 polypeptide B
621	5	1.2	16	21	AAY59333	Tyrosine phosphata
622	5	1.2	17	14	AAR31218	Immunosuppressive
623	5	1.2	17	19	AAW78547	SH2 domain binding
624	5	1.2	17	20	AAY59642	Secreted protein e
625	5	1.2	17	20	AAY25450	Human secreted pro
626	5	1.2	17	20	AAY35881	Signal peptide of
627	5	1.2	17	20	AAY05653	HIV-1 group M gp41
628	5	1.2	17	20	AAY12669	Human 5' EST secre
629	5	1.2	17	20	AAY11368	Human 5' EST secre
630	5	1.2	17	20	AAY11710	Peptide encoded by
631	5	1.2	17	20	AAY12515	Human 5' EST signa
632	5	1.2	17	20	AAY04165	Human secreted pro
633	5	1.2	17	20	AAY04147	Human secreted pro
634	5	1.2	17	20	AAW81991	Ehrlichia sp. 51 k
635	5	1.2	17	21	AAB30272	CD4+ T-cell activa
636	5	1.2	17	21	AAG00013	Human secreted pro
637	5	1.2	17	21	AAY76968	HIV-2 gp41 inhibit
638	5	1.2	17	21	AAY76420	Fragment of human
639	5	1.2	17	21	AAY64645	Human 48-19-3-G1-F
640	5	1.2	18	19	AAW54426	Human PS112 protei
641	5	1.2	18	21	AAB30273	CD4+ T-cell activa
642	5	1.2	18	21	AAB30274	CD4+ T-cell activa
643	5	1.2	18	21	AAB08416	Antigen peptide de
644	5	1.2	19	9	AAP80606	Sequence encoded b
645	5	1.2	19	13	AAR27825	Sequence of glycop
646	5	1.2	19	15	AAR46610	Malarial PfEMP3 19
647	5	1.2	19	17	AAR97973	Antigenic fragment
648	5	1.2	19	19	AAW37155	Mammalian Ena (Men
649	5	1.2	19	20	AAY42213	Oestrogen receptor
650	5	1.2	20	5	AAP40672	Sequence encoded b
651	5	1.2	20	11	AAR04979	Papilloma virus ty
652	5	1.2	20	12	AAR15560	Immunopeptide #11
653	5	1.2	20	14	AAR33481	Cytotoxic T-lympho
654	5	1.2	20	17	AAW08059	HIV peptide #44.
655	5	1.2	20	18	AAY26376	Erythropoietin rec
656	5	1.2	20	18	AAY13721	Erythropoietin rec
657	5	1.2	20	18	AAW27035	Monomer subunit of
658	5	1.2	20	19	AAW23879	HIV-1 strain MN pe
659	5	1.2	20	20	AAY39417	Human interleukin-
660	5	1.2	20	20	AAY29364	Original 3D profil
661	5	1.2	20	21	AAB35652	Mouse PAR-2 extrac
662	5	1.2	20	21	AAB45252	Human secreted pro
663	5	1.2	21	12	AAR11610	Flowering regulato
664	5	1.2	21	18	AAW15079	Lac21E monomer use
665	5	1.2	21	18	AAW31078	Mugwort pollen all
666	5	1.2	21	18	AAW00903	Human melanoma MAR
667	5	1.2	21	19	AAY21216	Human bcl2 proto-o
668	5	1.2	21	20	AAW95469	LS170 polypeptide

669	5	1.2	21	21	AAB30275	CD4+ T-cell activa
670	5	1.2	22	13	AAR25771	Anti-asthmatic pep
671	5	1.2	22	13	AAR25773	Anti-asthmatic pep
672	5	1.2	22	17	AAW00622	Plasmid palphaF-BZ
673	5	1.2	22	19	AAY20678	Human neurofilamen
674	5	1.2	22	20	AAW97829	Trichoderma reesei
675	5	1.2	23	13	AAR27003	Phaseolin signal s
676	5	1.2	24	7	AAP61516	Sequence encoded b
677	5	1.2	24	16	AAR65277	M2 region of volta
678	5	1.2	24	16	AAR65279	M2 region of volta
679	5	1.2	24	19	AAW48871	Pasteurella multoc
680	5	1.2	24	20	AAY02678	Human secreted pro
681	5	1.2	24	21	AAY97853	Staphylococcus aur
682	5	1.2	24	21	AAY97857	Staphylococcus aur
683	5	1.2	24	22	AAY72191	Peptide from regio
684	5	1.2	24	22	AAY72195	Peptide from regio
685	5	1.2	25	22	AAB64074	Human Clara cell p
686	5	1.2	25	22	AAB64076	Human Clara cell p
687	5	1.2	25	22	AAB64127	Human Clara cell p
688	5	1.2	25	22	AAB64128	Human Clara cell p
689	5	1.2	26	13	AAR27270	Signal peptide #4.
690	5	1.2	26	13	AAR24867	Sequence of peptid
691	5	1.2	26	20	AAY39260	G-protein coupled
692	5	1.2	26	21	AAY97351	Mammalian Bok Bcl-
693	5	1.2	26	22	AAB70377	BOK BH3 consensus
694	5	1.2	27	15	AAR63483	Rat MR77 serotonin
695	5	1.2	27	19	AAY20746	Human neurofilamen
696	5	1.2	27	20	AAY12859	Human 5' EST secre
697	5	1.2	27	20	AAW92293	Mouse alpha-1 (XVI
698	5	1.2	27	21	AAB37021	Bcl2 polypeptide B
699	5	1.2	27	21	AAB51843	Human secreted pro
700	5	1.2	27	21	AAY85645	Oxidoreductase int
701	5	1.2	27	21	AAB34128	Human secreted pro
702	5	1.2	28	13	AAR31584	GRF analogue #1.
703	5	1.2	28	17	AAW04494	Staphylococcal ent
704	5	1.2	28	17	AAR91398	ME#1 library deriv
705	5	1.2	28	17	AAR91384	ME#1 library deriv
706	5	1.2	28	18	AAW15085	Peptide monomer us
707	5	1.2	28	20	AAW73920	Staphylococcal ent
708	5	1.2	29	7	AAP61720	Fragment of sak ge
709	5	1.2	29	19	AAY85813	S. pneumoniae deri
710	5	1.2	29	20	AAY27704	Human secreted pro
711	5	1.2	29	20	AAY02684	Human secreted pro
712	5	1.2	29	20	AAW95466	LS170 polypeptide
713	5	1.2	29	21	AAB25600	Peptide encoded by
714	5	1.2	29	21	AAB25601	Human secreted pro
715	5	1.2	29	21	AAB25602	Human secreted pro
716	5	1.2	30	7	AAP61719	Fragment of sak ge
717	5	1.2	30	12	AAR10839	N-terminal of inte
718	5	1.2	30	18	AAW34148	Analogue #4 of HTL
719	5	1.2	30	20	AAY11534	Human 5' EST secre
720	5	1.2	30	20	AAW89959	Antigen 3 from clu
721	5	1.2	30	21	AAB51939	Human secreted pro
722	5	1.2	30	21	AAB27950	Human secreted pro
723	5	1.2	30	21	AAB39246	Human secreted pro
724	5	1.2	30	21	AAY89239	Core polypeptide f
725	5	1.2	30	22	AAB77592	Core polypeptide T
726	5	1.2	31	7	AAP61718	Fragment of sak ge
727	5	1.2	31	21	AAB51391	Human secreted pep
728	5	1.2	31	21	AAB39011	Human secreted pep
729	5	1.2	31	21	AAY55087	C. trachomatis maj
730	5	1.2	32	20	AAY11389	Human 5' EST secre
731	5	1.2	32	21	AAY44807	Calluna vulgaris h
732	5	1.2	32	21	AAY58650	Calluna vulgaris h
733	5	1.2	33	7	AAP61716	Fragment of sak ge
734	5	1.2	33	7	AAP61717	Fragment of sak ge
735	5	1.2	33	7	AAP61721	Fragment of sak ge
736	5	1.2	33	18	AAW00961	Heptad peptide 4he
737	5	1.2	33	21	AAB39136	Human secreted pro
738	5	1.2	33	21	AAB27822	Human secreted pro
739	5	1.2	33	21	AAB25611	Peptide encoded by
740	5	1.2	33	21	AAY65383	Human 5' EST relat
741	5	1.2	34	18	AAW28191	Staphylococcus aur
742	5	1.2	34	18	AAW15081	Lac28E monomer use
743	5	1.2	34	18	AAW17756	Peptide from phage
744	5	1.2	34	18	AAW18307	Peptide from phage
745	5	1.2	34	22	AAB77956	Core polypeptide T
746	5	1.2	35	2	AAP10428	CB7 type 24 strept
747	5	1.2	35	19	AAW79057	Rieske motif in Sy
748	5	1.2	35	19	AAW79029	Rieske motif in Sy
749	5	1.2	35	20	AAW96863	Nucleic acid bindi
750	5	1.2	35	20	AAW88674	Secreted protein e
751	5	1.2	35	21	AAB39335	Human secreted pro
752	5	1.2	35	21	AAB45311	Human secreted pro
753	5	1.2	35	21	AAB34306	Human secreted pro
754	5	1.2	35	21	AAB08962	Human secreted pro
755	5	1.2	35	21	AAB08969	Human secreted pro
756	5	1.2	35	21	AAY89552	Core polypeptide f
757	5	1.2	35	21	AAY89553	Core polypeptide f
758	5	1.2	35	21	AAY89554	Core polypeptide f
759	5	1.2	35	21	AAY89555	Core polypeptide f
760	5	1.2	35	21	AAY89556	Core polypeptide f
761	5	1.2	35	21	AAY89557	Core polypeptide f
762	5	1.2	35	21	AAY89558	Core polypeptide f
763	5	1.2	35	21	AAY89559	Core polypeptide f
764	5	1.2	35	21	AAY89560	Core polypeptide f
765	5	1.2	35	21	AAY89561	Core polypeptide f
766	5	1.2	35	21	AAY89562	Core polypeptide f
767	5	1.2	35	21	AAY59477	Metal oxide bindin
768	5	1.2	35	22	AAB77953	Core polypeptide T
769	5	1.2	35	22	AAB77954	Core polypeptide T
770	5	1.2	35	22	AAB77955	Core polypeptide T
771	5	1.2	35	22	AAB77957	Core polypeptide T
772	5	1.2	35	22	AAB77958	Core polypeptide T
773	5	1.2	35	22	AAB77959	Core polypeptide T
774	5	1.2	35	22	AAB77960	Core polypeptide T
775	5	1.2	35	22	AAB77961	Core polypeptide T
776	5	1.2	35	22	AAB77962	Core polypeptide T
777	5	1.2	35	22	AAB77963	Core polypeptide T
778	5	1.2	36	14	AAR32069	VLP reactive HPV16
779	5	1.2	36	17	AAW03554	Ligand scavenger r
780	5	1.2	36	21	AAB27802	Human secreted pro
781	5	1.2	36	21	AAB14539	HIV-1 isolate ADA
782	5	1.2	36	21	AAB14542	HIV-1 isolate JRFL
783	5	1.2	36	21	AAB14545	HIV-1 isolate 89.6
784	5	1.2	36	21	AAB14548	HIV-1 isolate BU91
785	5	1.2	36	21	AAB14553	HIV-1 isolate BZ16
786	5	1.2	36	21	AAB14558	HIV-1 isolate BE.V
787	5	1.2	36	21	AAB14561	HIV-1 isolate SE.S
788	5	1.2	36	22	AAB61440	Human TANGO 275 EG
789	5	1.2	37	17	AAW04313	Modified mouse obe
790	5	1.2	37	18	AAW10938	Polyclonal anti-fe
791	5	1.2	37	19	AAW43446	Scorpion toxin-rel
792	5	1.2	37	20	AAY02078	Kaliotoxin 2 pepti
793	5	1.2	37	21	AAB43726	Human cancer assoc
794	5	1.2	37	21	AAY64664	Human 5' EST relat
795	5	1.2	38	15	AAR55637	DP-118 - DP-107 an
796	5	1.2	38	15	AAR58458	TSAR binding domai
797	5	1.2	38	15	AAR58366	TSAR binding domai
798	5	1.2	38	16	AAR67703	DP-118 control pep
799	5	1.2	38	17	AAR98405	Peptide DP118, inh
800	5	1.2	38	17	AAR95524	Monoclonal antibod
801	5	1.2	38	18	AAW45040	Immunomodulatory p
802	5	1.2	38	18	AAW45043	Immunomodulatory p
803	5	1.2	38	18	AAW27629	Human immunodefici
804	5	1.2	38	18	AAW17018	DP-178-like peptid
805	5	1.2	38	20	AAY60463	Human normal bladd
806	5	1.2	38	20	AAY09490	Immunoactive pepti
807	5	1.2	38	20	AAY09487	Immunoactive pepti
808	5	1.2	38	21	AAB63074	Human secreted pro
809	5	1.2	38	21	AAB39336	Human secreted pro
810	5	1.2	38	21	AAB14540	HIV-1 isolate ADA
811	5	1.2	38	21	AAB14543	HIV-1 isolate JRFL
812	5	1.2	38	21	AAB14546	HIV-1 isolate 89.6
813	5	1.2	38	21	AAB14549	HIV-1 isolate BU91
814	5	1.2	38	21	AAB14551	HIV-1 isolate 92UG

815	5	1.2	38	21	AAB14554	HIV-1 isolate BZ16
816	5	1.2	38	21	AAB14556	HIV-1 isolate FLHH
817	5	1.2	38	21	AAB14559	HIV-1 isolate BE.V
818	5	1.2	38	21	AAB14562	HIV-1 isolate SE.S
819	5	1.2	38	21	AAG56794	Arabidopsis thalia
820	5	1.2	38	21	AAY89502	Core polypeptide f
821	5	1.2	38	21	AAY65379	Human 5' EST relat
822	5	1.2	38	22	AAB77903	Core polypeptide T
823	5	1.2	38	22	AAB54792	HIV antiviral acti
824	5	1.2	39	11	AAR06407	HTLV-1 correspondi
825	5	1.2	39	12	AAR13188	Peptide (XI) immun
826	5	1.2	39	17	AAR87555	Peptide #16 for th
827	5	1.2	39	17	AAR87541	Peptide #2 for the
828	5	1.2	39	20	AAY12631	Human 5' EST secre
829	5	1.2	39	20	AAY02729	Human secreted pro
830	5	1.2	39	21	AAB63080	Human secreted pro
831	5	1.2	39	21	AAB28015	Human secreted pro
832	5	1.2	39	21	AAG53225	Arabidopsis thalia
833	5	1.2	39	21	AAY56426	Helicobacter pylor
834	5	1.2	40	12	AAR13185	Peptide (V) immuno
835	5	1.2	40	17	AAR87540	Peptide #1 for the
836	5	1.2	40	18	AAW34143	HTLV gp46 fragment
837	5	1.2	40	18	AAW19086	Trypanosoma cruzi
838	5	1.2	40	19	AAW74750	Human secreted pro
839	5	1.2	40	20	AAY23282	Repeat sequence of
840	5	1.2	40	21	AAB26481	Synthetic peptide
841	5	1.2	40	21	AAB20686	Polymeric immunogl
842	5	1.2	40	21	AAY96989	PTH receptor left
843	5	1.2	40	21	AAG47155	Arabidopsis thalia
844	5	1.2	40	21	AAY44808	Protein encoded by
845	5	1.2	40	21	AAY45010	Calluna vulgaris h
846	5	1.2	40	21	AAY58651	Calluna vulgaris h
847	5	1.2	40	21	AAY55691	M. leprae ycfB pep
848	5	1.2	40	21	AAY64881	Human 5' EST relat
849	5	1.2	41	19	AAW37739	An N-terminal frag
850	5	1.2	41	19	AAW52171	Interferon-gamma i
851	5	1.2	41	21	AAB33901	Human secreted pro
852	5	1.2	42	18	AAW22049	Interferon gamma i
853	5	1.2	42	20	AAY12368	Human 5' EST secre
854	5	1.2	42	21	AAB16879	Bacteriophage Dp-1
855	5	1.2	42	21	AAG10311	Arabidopsis thalia
856	5	1.2	43	7	AAP61715	Fragment of sak ge
857	5	1.2	43	13	AAR31585	GRF analogue #2.
858	5	1.2	43	19	AAY21020	Human glial fibril
859	5	1.2	43	20	AAW88581	Secreted protein e
860	5	1.2	43	21	AAB33945	Human secreted pro
861	5	1.2	43	21	AAB34037	Human secreted pro
862	5	1.2	43	21	AAB34038	Human secreted pro
863	5	1.2	43	22	AAB69633	Targeting signal p
864	5	1.2	44	18	AAW28234	Amino acid sequenc
865	5	1.2	44	20	AAY19504	Amino acid sequenc
866	5	1.2	44	20	AAW74027	Human sucrose isom
867	5	1.2	44	21	AAB25691	Human secreted pro
868	5	1.2	44	21	AAG26640	Arabidopsis thalia
869	5	1.2	44	21	AAG55969	Arabidopsis thalia
870	5	1.2	44	21	AAG56386	Arabidopsis thalia
871	5	1.2	45	11	AAR07016	C-phycocyanin - la
872	5	1.2	45	19	AAY21005	Human glial fibril
873	5	1.2	45	20	AAY11750	Human 5' EST secre
874	5	1.2	45	20	AAY11567	Human 5' EST secre
875	5	1.2	45	21	AAB53857	Human colon cancer
876	5	1.2	45	21	AAB38529	Human secreted pro
877	5	1.2	45	21	AAB14538	HIV-1 isolate ADA
878	5	1.2	45	21	AAB14541	HIV-1 isolate JRFL
879	5	1.2	45	21	AAB14544	HIV-1 isolate 89.6
880	5	1.2	45	21	AAB14547	HIV-1 isolate BU91
881	5	1.2	45	21	AAB14550	HIV-1 isolate 92UG
882	5	1.2	45	21	AAB14552	HIV-1 isolate BZ16
883	5	1.2	45	21	AAB14555	HIV-1 isolate FLHH
884	5	1.2	45	21	AAB14557	HIV-1 isolate BE.V
885	5	1.2	45	21	AAB14560	HIV-1 isolate SE.S
886	5	1.2	45	22	AAB48246	Basic region of RO
887	5	1.2	46	20	AAY48438	Human prostate can
888	5	1.2	46	21	AAB28753	Protein fragment e
889	5	1.2	46	21	AAB34183	Gene 28 human secr
890	5	1.2	46	21	AAB06145	Caenorhabditis ele
891	5	1.2	46	22	AAB31652	Amino acid sequenc
892	5	1.2	47	21	AAB27175	Sendai virus parti
893	5	1.2	47	21	AAB51457	Human secreted pro
894	5	1.2	47	21	AAB51461	Human secreted pro
895	5	1.2	47	21	AAB32449	Human secreted pro
896	5	1.2	48	20	AAY13202	Human secreted pro
897	5	1.2	48	20	AAY11641	Human 5' EST secre
898	5	1.2	48	20	AAW83161	Rat orphan tyrosin
899	5	1.2	48	21	AAG24941	Arabidopsis thalia
900	5	1.2	48	22	AAB64551	Human secreted pro
901	5	1.2	49	8	AAP70036	Secretory signal s
902	5	1.2	49	21	AAG57102	Arabidopsis thalia
903	5	1.2	49	21	AAG60232	Arabidopsis thalia
904	5	1.2	49	22	AAB90567	Human secreted pro
905	5	1.2	50	20	AAY45264	Human secreted pro
906	5	1.2	50	20	AAY30888	Human secreted pro
907	5	1.2	50	21	AAY86292	Human secreted pro
908	5	1.2	50	22	AAB87119	Human TANGO 402 si
909	5	1.2	51	20	AAY33731	HIV peptide antige
910	5	1.2	51	21	AAB33929	Human secreted pro
911	5	1.2	51	21	AAB39100	Human secreted pro
912	5	1.2	51	21	AAG03447	Human secreted pro
913	5	1.2	51	21	AAY87215	Human secreted pro
914	5	1.2	52	6	AAP50255	Sequence of T-cell
915	5	1.2	52	7	AAP60075	Sequence of Hok pr
916	5	1.2	52	16	AAR70180	SPRU2 signal seque
917	5	1.2	52	20	AAY27655	Human secreted pro
918	5	1.2	52	21	AAY87154	Human secreted pro
919	5	1.2	53	18	AAW28033	Staphylococcus aur
920	5	1.2	53	19	AAW75396	Human brain glycog
921	5	1.2	53	20	AAY13042	Human secreted pro
922	5	1.2	53	21	AAB34706	Human secreted pro
923	5	1.2	53	21	AAG32840	Zea mays protein f
924	5	1.2	53	21	AAG40995	Zea mays protein f
925	5	1.2	53	21	AAG56729	Arabidopsis thalia
926	5	1.2	53	21	AAY50724	S. aureus sarA pro
927	5	1.2	53	22	AAB76582	Corynebacterium gl
928	5	1.2	54	7	AAP60428	Sequence encoded b
929	5	1.2	54	10	AAP91239	Epidermal growth f
930	5	1.2	54	10	AAP91238	Epidermal growth f
931	5	1.2	54	10	AAP91237	Epidermal growth f
932	5	1.2	54	10	AAP92078	Synthetic chimeric
933	5	1.2	54	10	AAP92079	Synthetic chimeric
934	5	1.2	54	13	AAR20077	Sequence encoded b
935	5	1.2	54	13	AAR22272	Human vitronectin
936	5	1.2	54	19	AAW79377	Staphylococcus aur
937	5	1.2	54	21	AAG57938	Zea mays protein f
938	5	1.2	54	21	AAG00362	Human secreted pro
939	5	1.2	54	21	AAG00997	Human secreted pro
940	5	1.2	55	20	AAY07911	Human secreted pro
941	5	1.2	55	21	AAB24577	Human secreted pro
942	5	1.2	55	21	AAG00332	Human secreted pro
943	5	1.2	55	21	AAY54577	A rat calcium/calm
944	5	1.2	56	20	AAY02885	Fragment of human
945	5	1.2	56	20	AAY01135	Secreted protein e
946	5	1.2	56	21	AAB45029	Human secreted pro
947	5	1.2	56	21	AAG60550	Arabidopsis thalia
948	5	1.2	56	21	AAG02566	Human secreted pro
949	5	1.2	56	21	AAY54578	A mouse calcium/ca
950	5	1.2	56	21	AAY86355	Human gene 8-encod
951	5	1.2	57	11	AAR08086	Feline T-cell lymp
952	5	1.2	57	16	AAR66178	Mouse T118 Flt3 li
953	5	1.2	57	17	AAR95435	Apomyoglobin ALPHA
954	5	1.2	57	20	AAY36261	Human secreted pro
955	5	1.2	57	20	AAW93838	Human 15S lipoxyge
956	5	1.2	57	21	AAG01773	Human secreted pro
957	5	1.2	58	19	AAY21183	Human bcl2 proto-o
958	5	1.2	58	20	AAY11565	Human 5' EST secre
959	5	1.2	58	21	AAB32894	Pinus radiata tran
960	5	1.2	58	21	AAG07657	Arabidopsis thalia

961	5	1.2	58	21	AAG61379	Arabidopsis thalia
962	5	1.2	58	21	AAG03325	Human secreted pro
963	5	1.2	58	21	AAY57375	Drosophila Shaker
964	5	1.2	58	21	AAY80566	Drosophila SHAKER1
965	5	1.2	59	17	AAW34246	SH3 domain of huma
966	5	1.2	59	17	AAW34229	SH3 domain C of mo
967	5	1.2	59	20	AAY38405	Human secreted pro
968	5	1.2	59	20	AAY27689	Human secreted pro
969	5	1.2	59	20	AAW88835	Polypeptide fragme
970	5	1.2	59	21	AAB11626	D. immitis ankyrin
971	5	1.2	59	21	AAG60085	Arabidopsis thalia
972	5	1.2	59	21	AAG60683	Arabidopsis thalia
973	5	1.2	59	21	AAG02198	Human secreted pro
974	5	1.2	59	21	AAY99584	Androctonus austra
975	5	1.2	59	22	AAB60789	Androctonus austra
976	5	1.2	60	20	AAY59924	Human myometrium t
977	5	1.2	60	20	AAY27648	Human secreted pro
978	5	1.2	60	21	AAB44703	Human secreted pro
979	5	1.2	60	21	AAG08060	Arabidopsis thalia
980	5	1.2	60	21	AAG25494	Arabidopsis thalia
981	5	1.2	60	21	AAG61623	Arabidopsis thalia
982	5	1.2	60	21	AAB15921	E. coli proliferat
983	5	1.2	60	21	AAG00540	Human secreted pro
984	5	1.2	60	21	AAY44806	Calluna vulgaris h
985	5	1.2	60	21	AAY58652	Calluna vulgaris h
986	5	1.2	60	22	AAB90654	Human secreted pro
987	5	1.2	61	14	AAR33732	Human intrinsic fa
988	5	1.2	61	18	AAW34665	Partial PSD-95 pro
989	5	1.2	61	19	AAY21104	Human bcl2 proto-o
990	5	1.2	61	21	AAG23761	Arabidopsis thalia
991	5	1.2	61	21	AAG40692	Zea mays protein f
992	5	1.2	61	21	AAG51828	Arabidopsis thalia
993	5	1.2	61	21	AAG60330	Arabidopsis thalia
994	5	1.2	62	14	AAR33391	Albumin-type prote
995	5	1.2	62	15	AAR53355	sCEC-SB37 fusion p
996	5	1.2	62	20	AAY76567	Human ovarian tumo
997	5	1.2	62	21	AAG10825	Arabidopsis thalia
998	5	1.2	62	21	AAG23760	Arabidopsis thalia
999	5	1.2	62	21	AAG48918	Arabidopsis thalia
1000	5	1.2	62	21	AAG51827	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAG28617
ID AAG28617 standard; Protein; 144 AA.
XX
AC AAG28617;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33903.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.9%; Score 8; DB 21; Length 144;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 378 SGPSLALS 385
||||||||
Db 74 sgpslals 81

RESULT 2
AAW72112
ID AAW72112 standard; Protein; 284 AA.
XX
AC AAW72112;
XX
DT 18-DEC-1998 (first entry)
XX
DE HSV-2 strain SB5 Contig ID 13 ORF#1 protein.
XX
KW HSV-2 strain SB5; immunological response induction; therapy;
KW antiviral identification; viral protein inhibitor.
XX
OS Herpes simplex virus type 2.
XX
FH Key Location/Qualifiers
FT Misc-difference 279..284
FT /note= "encoded by GG"
FT Misc-difference 280..284
FT /note= "unspecified amino acids"
XX
PN WO9820016-A1.
XX
PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97WO-US20016.
XX
PR 09-JUN-1997; 97US-0049018.
PR 04-NOV-1996; 96US-0030279.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
PI Esser KM, Leary JJ;
XX

DR WPI; 1998-286847/25.
DR N-PSDB; AAV62157.
XX
PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention
PT and treatment of infection or inducing immunological response in
PT mammal
XX
PS Claim 10; Page 86; 748pp; English.
XX
CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
CC sequence of the invention. This sequence was isolated from a HSV-2 strain
CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 13.
CC The proteins can be used for the treatment or prevention of disease, to
CC induce an immunological response in a mammal or to identify inhibitors,
CC activators or novel antivirals. Antagonists of the proteins can be used
CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
CC it can also be used to induce an immunological response in a mammal.
XX
SQ Sequence 284 AA;

Query Match 1.9%; Score 8; DB 19; Length 284;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 LKEELARF 114
||||||||
Db 141 lkeelarf 148

RESULT 3
AAW72195
ID AAW72195 standard; Protein; 785 AA.
XX
AC AAW72195;
XX
DT 13-JAN-1999 (first entry)
XX
DE HSV-2 strain SB5 Contig ID 15 ORF#30 protein.
XX
KW HSV-2 strain SB5; immunological response induction; therapy;
KW antiviral identification; viral protein inhibitor.
XX
OS Herpes simplex virus type 2.
XX
PN WO9820016-A1.
XX
PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97WO-US20016.
XX
PR 09-JUN-1997; 97US-0049018.
PR 04-NOV-1996; 96US-0030279.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
PI Esser KM, Leary JJ;
XX
DR WPI; 1998-286847/25.
DR N-PSDB; AAV62176.
XX
PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention
PT and treatment of infection or inducing immunological response in
PT mammal
XX
PS Claim 10; Page 122; 748pp; English.
XX
CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
CC sequence of the invention. This sequence was isolated from a HSV-2 strain
CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 15.
CC The proteins can be used for the treatment or prevention of disease, to
CC induce an immunological response in a mammal or to identify inhibitors,
CC activators or novel antivirals. Antagonists of the proteins can be used
CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
CC it can also be used to induce an immunological response in a mammal.
XX
SQ Sequence 785 AA;

Query Match 1.9%; Score 8; DB 19; Length 785;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 LKEELARF 114
||||||||
Db 640 lkeelarf 647

RESULT 4
AAG47201
ID AAG47201 standard; Protein; 3636 AA.
XX
AC AAG47201;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 59465.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```
  Query Match             1.9%;  Score 8;  DB 21;  Length 3636;
  Best Local Similarity   100.0%;  Pred. No. 75;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     307 KVLTSLNR 314
           ||||||||
Db    1210 kvltslnr 1217


RESULT   5
AAG47200
ID   AAG47200 standard; Protein; 3712 AA.
XX
AC   AAG47200;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 59464.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
```

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.9%; Score 8; DB 21; Length 3712;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     307 KVLTSLNR 314
           ||||||||
Db    1286 kvltslnr 1293
```

RESULT 6
AAG47199
ID AAG47199 standard; Protein; 3738 AA.
XX
AC AAG47199;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 59463.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.9%; Score 8; DB 21; Length 3738;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 307 KVLTSLNR 314
||||||||
Db 1312 kvltslnr 1319

RESULT 7
AAB72418
ID AAB72418 standard; peptide; 13 AA.
XX
AC AAB72418;

XX
DT 08-MAY-2001 (first entry)
XX
DE Peptide used in the characterisation of MBL ligands.
XX
KW Lectin complement pathway; complement activation; mannose binding lectin;
KW cellular injury.
XX
OS Cytisus sessilifolius.
XX
FH Key Location/Qualifiers
FT Misc-difference 13
FT /label= Arg, Lys
XX
PN WO200112212-A1.
XX
PD 22-FEB-2001.
XX
PF 14-AUG-2000; 2000WO-US22123.
XX
PR 13-AUG-1999; 99US-0148815.
XX
PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI Stahl GL, Lekowski R;
XX
DR WPI; 2001-218308/22.
XX
PT Inhibiting lectin complement pathway associated complement activation,
PT useful for inhibiting cellular injury, comprises contacting mammalian
PT cell with surface exposed mannose binding lectin ligand (MBL) with MBL
PT receptor antagonist -
XX
PS Example 2; Page 43; 87pp; English.
XX
CC The present invention relates to a method for inhibiting lectin
CC complement pathway (LCP) associated complement activation. The method
CC comprising contacting a mammalian cell with surface exposed mannose
CC binding lectin (MBL) ligand with an MBL receptor antagonist (RA) to
CC inhibit LCP-associated complement activation. LCP activation normally
CC occurs through the binding of MBL and its two associated serine
CC proteases, MASP-1 and MASP-2, to carbohydrates on the surface of
CC microorganisms. A composition comprising an isolated RA is useful for the
CC treatment of an MBL mediated disorder, where RA is a legume derived
CC lectin or an isolated peptide fragment of a legume derived lectin.
CC The composition of the present invention is useful for inhibiting a
CC cellular injury mediated by LCP associated complement activation which
CC contributes to tissue injury associated with atherosclerosis, pulmonary
CC system, arthritis, myocardial infarction, ischaemia, reperfusion,
CC transplantation, CPB, stroke, adult respiratory distress syndrome (ARDS),
CC lupus, systemic lupus erythematosus and dialysis. The present sequence is
CC a peptide which was used in the present invention in the characterisation
CC of MBL ligands.
XX
SQ Sequence 13 AA;

Query Match 1.6%; Score 7; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 GLKEELA 112
|||||||
Db 4 glkeela 10

RESULT 8
AAB38359
ID AAB38359 standard; Protein; 97 AA.
XX
AC AAB38359;
XX
DT 31-JAN-2001 (first entry)
XX
DE Human secreted protein encoded by gene 39 clone HTLIQ05.
XX
KW Immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
KW cytostatic; cardiant; vasotropic; cerebroprotective; neuroprotective;
KW nootropic; antibacterial; virucide; fungicide; opthalmalogical; human;
KW vulnerary; gene therapy; infection; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200061623-A1.
XX
PD 19-OCT-2000.
XX
PF 05-APR-2000; 2000WO-US08979.
XX
PR 09-APR-1999; 99US-0128693.
PR 26-APR-1999; 99US-0130991.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Ni J, Komatsoulis GA, Rosen CA, Soppet DR, Shi Y;
PI Lafleur DW, Olsen HS, Ebner R, Florence KA, Moore PA, Birse CE;
PI Young PE;
XX
DR WPI; 2000-647418/62.
XX
PT New nucleic acid molecules encoding 62 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
PS Claim 11; Page 628-629; 716pp; English.
XX
CC Sequences AAB38321-B38396 represent the amino acid sequences of 62
CC human secreted proteins encoded by the genes AAC69512-C69587. The genes
CC and proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated from
CC a range of human tissues disclosed in the specification. The nucleic
CC acids, proteins, antibodies and (ant)agonists are useful in the
CC diagnosis, treatment and prevention of: (a) autoimmune diseases e.g.
CC rheumatoid arthritis; (b) hyperproliferative disorders e.g. neoplasms
CC of the breast or liver; (c) cardiovascular disorders e.g. cardiac
CC arrest; (d) cerebrovascular disorders e.g. cerebral ischemia; (e)
CC angiogenesis; (f) nervous system disorders e.g. Alzheimer's disease; (g)
CC infections caused by bacteria, viruses and fungi; and (h) ocular
CC disorders e.g. corneal infection. The polypeptides can also be used to
CC aid wound healing and epithelial cell proliferation, to prevent skin
CC aging due to sunburn, to maintain organs before transplantation, for
CC supporting cell culture of primary tissues, to regenerate tissues and in
CC chemotaxis.
XX
SQ Sequence 97 AA;

Query Match 1.6%; Score 7; DB 21; Length 97;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 LKEELAR 113
|||||||
Db 43 lkeelar 49

RESULT 9
AAB80170
ID AAB80170 standard; Protein; 138 AA.
XX
AC AAB80170;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:1074.
XX

KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100843-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB00923.
XX
PR 25-JUN-1999; 99US-0141031.
PR 01-JUL-1999; 99DE-1030476.
PR 02-JUL-1999; 99US-0142101.
PR 08-JUL-1999; 99DE-1031415.
PR 08-JUL-1999; 99DE-1031418.
PR 08-JUL-1999; 99DE-1031419.
PR 08-JUL-1999; 99DE-1031420.
PR 08-JUL-1999; 99DE-1031424.
PR 08-JUL-1999; 99DE-1031428.
PR 08-JUL-1999; 99DE-1031434.
PR 08-JUL-1999; 99DE-1031435.
PR 08-JUL-1999; 99DE-1031443.
PR 08-JUL-1999; 99DE-1031453.
PR 08-JUL-1999; 99DE-1031457.
PR 08-JUL-1999; 99DE-1031465.
PR 08-JUL-1999; 99DE-1031478.
PR 08-JUL-1999; 99DE-1031510.
PR 08-JUL-1999; 99DE-1031541.
PR 08-JUL-1999; 99DE-1031573.
PR 08-JUL-1999; 99DE-1031592.
PR 08-JUL-1999; 99DE-1031632.
PR 08-JUL-1999; 99DE-1031634.
PR 08-JUL-1999; 99DE-1031636.
PR 09-JUL-1999; 99DE-1032125.
PR 09-JUL-1999; 99DE-1032126.
PR 09-JUL-1999; 99DE-1032130.
PR 09-JUL-1999; 99DE-1032186.
PR 09-JUL-1999; 99DE-1032206.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032228.
PR 09-JUL-1999; 99DE-1032229.
PR 09-JUL-1999; 99DE-1032230.
PR 14-JUL-1999; 99DE-1032922.
PR 14-JUL-1999; 99DE-1032926.
PR 14-JUL-1999; 99DE-1032928.
PR 14-JUL-1999; 99DE-1033004.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 12-AUG-1999; 99US-0148613.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040832.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041380.
PR 31-AUG-1999; 99DE-1041394.
PR 31-AUG-1999; 99DE-1041396.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042124.
PR 03-SEP-1999; 99DE-1042129.
PR 09-MAR-2000; 2000US-0187970.
XX
PA (BADI) BASF AG.
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR WPI; 2001-137957/14.
DR N-PSDB; AAF72289.
XX
PT Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in
PT microorganisms, including organic acids, nonproteinogenic amino acids,
PT and purine and pyrimidine bases -
XX
PS Claim 20; Page 1628; 1737pp; English.
XX
CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX
SQ Sequence 138 AA;

Query Match 1.6%; Score 7; DB 22; Length 138;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     320 RFLRYNA 326
           |||||||
Db      71 rflryna 77
```

RESULT 10
AAB44011
ID AAB44011 standard; Protein; 143 AA.
XX
AC AAB44011;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated protein sequence SEQ ID NO:1456.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05882.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587533/55.
DR N-PSDB; AAC78220.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides

PT useful for treating or diagnosing e.g. cancer -
XX
PS Claim 11; Page 2136; 2352pp; English.
XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnerary; immunomodulator;
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coaguiant;
CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.
XX
SQ Sequence 143 AA;

Query Match 1.6%; Score 7; DB 21; Length 143;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 GLKEELA 112
|||||||
Db 82 glkeela 88

RESULT 11
AAY58586
ID AAY58586 standard; Protein; 145 AA.
XX
AC AAY58586;
XX
DT 10-APR-2000 (first entry)
XX
DE Sorangium cellulosum protein Orf 7.
XX
KW Orf 5; epothilone biosynthesis; taxol substitute; anticancer.
XX
OS Sorangium cellulosum.
XX
PN WO9966028-A2.
XX
PD 23-DEC-1999.
XX
PF 16-JUN-1999; 99WO-EP04171.
XX
PR 18-JUN-1998; 98US-0099504.
PR 24-SEP-1998; 98US-0101631.
PR 05-FEB-1999; 99US-0118906.
XX
PA (NOVS) NOVARTIS AG.
PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI Schupp T, Ligon JM, Molnar I, Zirkle R, Goerlach J, Cyr D;
XX
DR WPI; 2000-097741/08.
DR N-PSDB; AAZ55887.
XX
PT New isolated epothilone synthase genes, used for the recombinant
PT production of epothilone for use in cancer therapy -
XX
PS Example 12; Page 166-167; 174pp; English.
XX
CC The invention relates to nucleic acid sequences encoding epothilone
CC biosynthetic enzymes from Sorangium cellulosum. Epothilones A and B are
CC 16-membered macrocyclic polyketides with an acylcysteine-derived
CC starter unit; polyketides being synthesised from two-carbon building
CC blocks, the beta-carbon of which always carries a keto group. Each round
CC of two-carbon addition is carried out by a complex of enzymes known as
CC the polyketide synthase in a manner similar to fatty acid biosynthesis.
CC EPOS A (AAY58573) and EPOS P (AAY58574) are involved in formation of
CC the thiazole ring formation of epothilones, and EPOS B, EPOS C, EPOS D
CC and EPOS E (AAY58575-Y58578) are involved in polyketide backbone
CC formation. EPO F (AAY58579) is an epothilone macrolactone oxidase, and
CC the proteins Orf 3 (AAY58582) and Orf14 (AAY58593) are thought to be
CC involved in transport. Epothilones mimic the biological activity of
CC taxol, and may be substituted for taxol in cancer chemotherapeutic
CC compositions. Epothilones exhibit a much lower drop in potency against a
CC multiply drug-resistant cell line compared with taxol, and are
CC considerably less efficiently exported from such cells by the multidrug
CC resistance protein (MDR, or P-glycoprotein). Despite the potential of
CC epothilones as anticancer agents, they are problematical to produce on a
CC large scale. Epothilones are too complex for industrial scale chemical
CC synthesis, and Sorangium cellulosum is difficult to ferment, producing
CC poor yields of epothilones. The nucleic acids of the invention may be
CC used for the recombinant production of epothilones in a heterologous host
CC that is more amenable to fermentation. Sequences AAY58580,
CC AAY58583-Y58592 and AAY58594 represent proteins encoded by open reading
CC frames (ORFs) of a Sorangium cellulosum 68.75 kb contig which covers the
CC epothilone biosynthesis genes. The present sequence has no homologues in
CC the sequence databases.
XX
SQ Sequence 145 AA;

Query Match 1.6%; Score 7; DB 21; Length 145;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 SVRPCIR 89
|||||||
Db 127 svrpcir 133

RESULT 12
AAG19392
ID AAG19392 standard; Protein; 179 AA.
XX
AC AAG19392;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 21175.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.

PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 7; DB 21; Length 179;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    378 SGPSLAL 384
          |||||||
Db     93 sgpslal 99
```

RESULT 13
AAG32444
ID AAG32444 standard; Protein; 179 AA.
XX
AC AAG32444;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 39139.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 7; DB 21; Length 179;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    378 SGPSLAL 384
          |||||||
Db     93 sgpslal 99
```

RESULT 14
AAG47917
ID AAG47917 standard; Protein; 179 AA.
XX
AC AAG47917;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 60448.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 7; DB 21; Length 179;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 378 SGPSLAL 384
|||||||
Db 93 sgpslal 99

RESULT 15
AAG19391
ID AAG19391 standard; Protein; 180 AA.
XX
AC AAG19391;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 21174.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 7; DB 21; Length 180;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 378 SGPSLAL 384
|||||||
Db 94 sgpslal 100

Search completed: October 30, 2001, 15:29:50
Job time: 476 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 15:25:09 ; Search time 13.4 Seconds
(without alignments)
654.589 Million cell updates/sec

Title: US-09-619-049-855
Perfect score: 426
Sequence: 1 MRLPWVIFCTVLLLIFTNNS.....NKFTNEEEGRINCINFDSQT 426

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7	1.6	145	3	US-09-335-409-15	Sequence 15, Appl
2	7	1.6	385	3	US-09-053-866-2	Sequence 2, Appli
3	7	1.6	400	1	US-07-730-953-2	Sequence 2, Appli
4	7	1.6	741	4	US-09-001-984C-106	Sequence 106, App
5	7	1.6	1261	4	US-09-208-742-4	Sequence 4, Appli
6	6	1.4	22	1	US-08-447-411-71	Sequence 71, Appl
7	6	1.4	22	2	US-08-662-227-29	Sequence 29, Appl
8	6	1.4	102	2	US-08-480-473B-47	Sequence 47, Appl
9	6	1.4	102	3	US-08-915-213-47	Sequence 47, Appl
10	6	1.4	102	4	US-09-235-217-47	Sequence 47, Appl
11	6	1.4	131	2	US-08-466-583-5	Sequence 5, Appli
12	6	1.4	131	5	PCT-US95-07820-5	Sequence 5, Appli
13	6	1.4	146	1	US-08-240-712-21	Sequence 21, Appl
14	6	1.4	146	1	US-08-240-712-22	Sequence 22, Appl
15	6	1.4	146	1	US-08-240-712-23	Sequence 23, Appl
16	6	1.4	146	1	US-08-443-890-21	Sequence 21, Appl
17	6	1.4	146	1	US-08-443-890-22	Sequence 22, Appl
18	6	1.4	146	1	US-08-443-890-23	Sequence 23, Appl
19	6	1.4	146	2	US-08-484-686B-67	Sequence 67, Appl
20	6	1.4	146	4	US-08-463-160B-68	Sequence 68, Appl
21	6	1.4	146	5	PCT-US92-09752-21	Sequence 21, Appl
22	6	1.4	146	5	PCT-US92-09752-22	Sequence 22, Appl
23	6	1.4	146	5	PCT-US92-09752-23	Sequence 23, Appl
24	6	1.4	147	4	US-09-058-562-21	Sequence 21, Appl
25	6	1.4	151	3	US-08-584-031-11	Sequence 11, Appl
26	6	1.4	179	3	US-08-911-321-3	Sequence 3, Appli
27	6	1.4	193	1	US-08-106-507-2	Sequence 2, Appli
28	6	1.4	213	2	US-08-916-770-1	Sequence 1, Appli
29	6	1.4	213	2	US-09-102-593-1	Sequence 1, Appli
30	6	1.4	213	2	US-08-437-013-2	Sequence 2, Appli
31	6	1.4	216	1	US-08-106-507-10	Sequence 10, Appl
32	6	1.4	216	1	US-08-446-922-8	Sequence 8, Appli
33	6	1.4	216	5	PCT-US93-10034-8	Sequence 8, Appli
34	6	1.4	263	4	US-08-845-258-18	Sequence 18, Appl
35	6	1.4	263	4	US-08-990-571-18	Sequence 18, Appl
36	6	1.4	267	4	US-08-845-258-28	Sequence 28, Appl
37	6	1.4	267	4	US-08-990-571-28	Sequence 28, Appl
38	6	1.4	271	4	US-09-077-675A-12	Sequence 12, Appl
39	6	1.4	275	1	US-07-611-528A-2	Sequence 2, Appli
40	6	1.4	275	1	US-08-083-946-2	Sequence 2, Appli
41	6	1.4	275	3	US-08-452-915-2	Sequence 2, Appli
42	6	1.4	291	4	US-09-328-869-2	Sequence 2, Appli
43	6	1.4	294	4	US-08-845-258-46	Sequence 46, Appl
44	6	1.4	294	4	US-08-990-571-46	Sequence 46, Appl
45	6	1.4	295	1	US-08-118-270-79	Sequence 79, Appl
46	6	1.4	295	5	PCT-US93-08528-79	Sequence 79, Appl
47	6	1.4	302	4	US-09-077-675A-7	Sequence 7, Appli
48	6	1.4	303	4	US-08-845-258-23	Sequence 23, Appl
49	6	1.4	303	4	US-08-990-571-23	Sequence 23, Appl
50	6	1.4	310	4	US-08-845-258-19	Sequence 19, Appl
51	6	1.4	310	4	US-08-990-571-19	Sequence 19, Appl
52	6	1.4	318	1	US-08-349-696-25	Sequence 25, Appl
53	6	1.4	318	1	US-08-233-009-25	Sequence 25, Appl
54	6	1.4	318	2	US-08-560-231-25	Sequence 25, Appl
55	6	1.4	318	4	US-09-080-704A-25	Sequence 25, Appl
56	6	1.4	320	4	US-09-092-437-2	Sequence 2, Appli
57	6	1.4	322	2	US-08-428-414A-2	Sequence 2, Appli
58	6	1.4	324	2	US-08-813-940-2	Sequence 2, Appli
59	6	1.4	330	4	US-09-188-930-144	Sequence 144, App
60	6	1.4	330	4	US-09-188-930-278	Sequence 278, App
61	6	1.4	358	2	US-08-465-971B-4	Sequence 4, Appli
62	6	1.4	361	4	US-09-077-675A-8	Sequence 8, Appli
63	6	1.4	364	2	US-08-651-940-2	Sequence 2, Appli
64	6	1.4	364	4	US-09-025-580-37	Sequence 37, Appl
65	6	1.4	364	4	US-09-077-675A-16	Sequence 16, Appl
66	6	1.4	366	4	US-09-077-675A-13	Sequence 13, Appl
67	6	1.4	367	4	US-08-845-258-20	Sequence 20, Appl
68	6	1.4	367	4	US-08-845-258-49	Sequence 49, Appl
69	6	1.4	367	4	US-08-990-571-20	Sequence 20, Appl
70	6	1.4	367	4	US-08-990-571-49	Sequence 49, Appl
71	6	1.4	372	4	US-09-286-904-24	Sequence 24, Appl
72	6	1.4	373	2	US-08-480-473B-3	Sequence 3, Appli
73	6	1.4	373	3	US-08-915-213-3	Sequence 3, Appli
74	6	1.4	373	4	US-09-235-217-3	Sequence 3, Appli
75	6	1.4	373	5	PCT-US96-10251-3	Sequence 3, Appli
76	6	1.4	414	2	US-08-984-171-1	Sequence 1, Appli
77	6	1.4	432	2	US-08-705-660-18	Sequence 18, Appl
78	6	1.4	432	3	US-08-989-045-18	Sequence 18, Appl
79	6	1.4	439	3	US-08-993-359-24	Sequence 24, Appl
80	6	1.4	439	3	US-09-221-654-2	Sequence 2, Appli
81	6	1.4	439	3	US-08-989-358A-2	Sequence 2, Appli
82	6	1.4	450	1	US-08-194-338-5	Sequence 5, Appli
83	6	1.4	450	1	US-08-196-989B-11	Sequence 11, Appl
84	6	1.4	450	1	US-08-444-734A-8	Sequence 8, Appli
85	6	1.4	450	2	US-08-760-936-11	Sequence 11, Appl
86	6	1.4	464	1	US-08-252-995D-6	Sequence 6, Appli
87	6	1.4	464	2	US-08-834-108-6	Sequence 6, Appli
88	6	1.4	471	2	US-08-466-583-2	Sequence 2, Appli
89	6	1.4	471	5	PCT-US95-07820-2	Sequence 2, Appli
90	6	1.4	477	1	US-08-191-338A-2	Sequence 2, Appli
91	6	1.4	486	2	US-08-870-518-3	Sequence 3, Appli
92	6	1.4	489	4	US-08-942-012B-31	Sequence 31, Appl
93	6	1.4	501	1	US-08-331-394-4	Sequence 4, Appli
94	6	1.4	501	1	US-08-250-858-4	Sequence 4, Appli
95	6	1.4	501	1	US-08-446-915-4	Sequence 4, Appli
96	6	1.4	501	2	US-08-744-139-4	Sequence 4, Appli
97	6	1.4	501	5	PCT-US95-06639-4	Sequence 4, Appli
98	6	1.4	529	3	US-09-080-044-2	Sequence 2, Appli
99	6	1.4	532	1	US-08-288-405A-10	Sequence 10, Appl
100	6	1.4	571	4	US-08-809-326A-6	Sequence 6, Appli

101	6	1.4	591	4	US-09-352-159-44	Sequence 44, Appl
102	6	1.4	591	4	US-09-352-159-46	Sequence 46, Appl
103	6	1.4	598	4	US-09-352-159-40	Sequence 40, Appl
104	6	1.4	598	4	US-09-352-159-42	Sequence 42, Appl
105	6	1.4	600	4	US-09-352-159-23	Sequence 23, Appl
106	6	1.4	600	4	US-09-352-159-36	Sequence 36, Appl
107	6	1.4	600	4	US-09-352-159-38	Sequence 38, Appl
108	6	1.4	600	4	US-09-352-168-23	Sequence 23, Appl
109	6	1.4	622	4	US-09-305-381-2	Sequence 2, Appli
110	6	1.4	632	1	US-08-295-814E-10	Sequence 10, Appl
111	6	1.4	632	4	US-09-343-361-10	Sequence 10, Appl
112	6	1.4	632	5	PCT-US93-01959-10	Sequence 10, Appl
113	6	1.4	634	1	US-07-779-049-3	Sequence 3, Appli
114	6	1.4	634	1	US-08-080-240-3	Sequence 3, Appli
115	6	1.4	727	4	US-08-989-385-4	Sequence 4, Appli
116	6	1.4	739	2	US-08-836-943-2	Sequence 2, Appli
117	6	1.4	770	1	US-08-525-654A-1	Sequence 1, Appli
118	6	1.4	771	1	US-08-525-654A-3	Sequence 3, Appli
119	6	1.4	788	1	US-08-194-338-12	Sequence 12, Appl
120	6	1.4	805	2	US-08-480-473B-4	Sequence 4, Appli
121	6	1.4	805	3	US-08-915-213-4	Sequence 4, Appli
122	6	1.4	805	4	US-09-235-217-4	Sequence 4, Appli
123	6	1.4	805	5	PCT-US96-10251-4	Sequence 4, Appli
124	6	1.4	826	1	US-08-785-241-6	Sequence 6, Appli
125	6	1.4	826	2	US-08-480-473B-2	Sequence 2, Appli
126	6	1.4	826	3	US-08-915-213-2	Sequence 2, Appli
127	6	1.4	826	3	US-09-148-547-2	Sequence 2, Appli
128	6	1.4	826	4	US-09-235-217-2	Sequence 2, Appli
129	6	1.4	826	5	PCT-US96-10251-2	Sequence 2, Appli
130	6	1.4	828	2	US-08-993-228-21	Sequence 21, Appl
131	6	1.4	829	4	US-09-352-159-33	Sequence 33, Appl
132	6	1.4	829	4	US-09-352-168-33	Sequence 33, Appl
133	6	1.4	855	2	US-08-482-090-12	Sequence 12, Appl
134	6	1.4	856	2	US-08-481-700B-8	Sequence 8, Appli
135	6	1.4	856	2	US-09-007-383-16	Sequence 16, Appl
136	6	1.4	906	1	US-08-254-573-2	Sequence 2, Appli
137	6	1.4	906	1	US-08-687-379-2	Sequence 2, Appli
138	6	1.4	906	1	US-08-687-379-4	Sequence 4, Appli
139	6	1.4	907	1	US-07-718-575-2	Sequence 2, Appli
140	6	1.4	907	1	US-08-481-206-2	Sequence 2, Appli
141	6	1.4	907	2	US-08-486-269A-2	Sequence 2, Appli
142	6	1.4	915	2	US-08-480-917-2	Sequence 2, Appli
143	6	1.4	926	4	US-08-855-910-2	Sequence 2, Appli
144	6	1.4	974	3	US-08-938-291A-4	Sequence 4, Appli
145	6	1.4	1012	1	US-08-219-262B-10	Sequence 10, Appl
146	6	1.4	1012	3	US-09-031-655-10	Sequence 10, Appl
147	6	1.4	1013	2	US-08-708-541A-34	Sequence 34, Appl
148	6	1.4	1127	4	US-09-150-460B-11	Sequence 11, Appl
149	6	1.4	1349	3	US-08-938-291A-6	Sequence 6, Appli
150	6	1.4	1494	3	US-08-755-587-186	Sequence 186, App
151	6	1.4	1536	4	US-09-413-814-10	Sequence 10, Appl
152	6	1.4	1642	1	US-08-447-411-45	Sequence 45, Appl
153	6	1.4	1642	2	US-08-662-227-2	Sequence 2, Appli
154	6	1.4	2183	1	US-08-348-891A-7	Sequence 7, Appli
155	6	1.4	2183	2	US-08-905-817-7	Sequence 7, Appli
156	6	1.4	2471	3	US-09-112-450-4	Sequence 4, Appli
157	5	1.2	5	2	US-08-747-137-129	Sequence 129, App
158	5	1.2	5	4	US-09-177-249-134	Sequence 134, App
159	5	1.2	5	6	5217869-12	Patent No. 5217869
160	5	1.2	7	1	US-08-182-175A-69	Sequence 69, Appl
161	5	1.2	7	1	US-08-182-175A-72	Sequence 72, Appl
162	5	1.2	7	1	US-08-182-175A-75	Sequence 75, Appl
163	5	1.2	7	1	US-08-346-333-28	Sequence 28, Appl
164	5	1.2	7	2	US-08-452-724A-48	Sequence 48, Appl
165	5	1.2	7	5	PCT-US91-07506-28	Sequence 28, Appl
166	5	1.2	7	5	PCT-US92-06412-69	Sequence 69, Appl
167	5	1.2	7	5	PCT-US92-06412-72	Sequence 72, Appl
168	5	1.2	7	5	PCT-US92-06412-75	Sequence 75, Appl
169	5	1.2	8	1	US-08-266-514-9	Sequence 9, Appli
170	5	1.2	8	2	US-08-654-604-9	Sequence 9, Appli
171	5	1.2	8	2	US-08-310-912A-140	Sequence 140, App
172	5	1.2	8	5	PCT-US95-04589-140	Sequence 140, App
173	5	1.2	9	2	US-08-417-174-10	Sequence 10, Appl
174	5	1.2	9	2	US-08-231-565A-10	Sequence 10, Appl
175	5	1.2	9	2	US-08-637-759B-471	Sequence 471, App
176	5	1.2	9	2	US-08-318-856A-5	Sequence 5, Appli
177	5	1.2	9	2	US-09-007-961-10	Sequence 10, Appl
178	5	1.2	9	3	US-08-871-355A-471	Sequence 471, App
179	5	1.2	9	4	US-09-314-742-6	Sequence 6, Appli
180	5	1.2	9	4	US-09-551-510-6	Sequence 6, Appli
181	5	1.2	9	4	US-09-551-510-7	Sequence 7, Appli
182	5	1.2	10	1	US-08-077-797A-24	Sequence 24, Appl
183	5	1.2	10	2	US-08-417-174-22	Sequence 22, Appl
184	5	1.2	10	2	US-08-231-565A-22	Sequence 22, Appl
185	5	1.2	10	2	US-09-007-961-22	Sequence 22, Appl
186	5	1.2	10	5	PCT-US92-07218-3	Sequence 3, Appli
187	5	1.2	10	5	PCT-US94-01238-24	Sequence 24, Appl
188	5	1.2	11	1	US-08-432-694-22	Sequence 22, Appl
189	5	1.2	11	3	US-09-008-481A-11	Sequence 11, Appl
190	5	1.2	12	2	US-08-752-852A-158	Sequence 158, App
191	5	1.2	12	2	US-08-752-852A-171	Sequence 171, App
192	5	1.2	13	1	US-08-082-844-3	Sequence 3, Appli
193	5	1.2	13	1	US-08-191-866D-33	Sequence 33, Appl
194	5	1.2	13	2	US-08-185-949B-33	Sequence 33, Appl
195	5	1.2	13	2	US-08-752-852A-176	Sequence 176, App
196	5	1.2	13	3	US-08-752-853-44	Sequence 44, Appl
197	5	1.2	13	4	US-09-206-059-13	Sequence 13, Appl
198	5	1.2	13	5	PCT-US94-07040-3	Sequence 3, Appli
199	5	1.2	14	1	US-07-837-892-1	Sequence 1, Appli
200	5	1.2	14	2	US-08-637-759B-250	Sequence 250, App
201	5	1.2	14	3	US-08-871-355A-250	Sequence 250, App
202	5	1.2	14	4	US-09-155-427-14	Sequence 14, Appl
203	5	1.2	15	1	US-07-837-892-2	Sequence 2, Appli
204	5	1.2	15	4	US-08-602-999A-330	Sequence 330, App
205	5	1.2	15	5	PCT-US93-06751-8	Sequence 8, Appli
206	5	1.2	16	1	US-07-942-245-158	Sequence 158, App
207	5	1.2	16	4	US-09-323-433A-33	Sequence 33, Appl
208	5	1.2	16	4	US-09-323-433A-41	Sequence 41, Appl
209	5	1.2	16	4	US-08-669-304-10	Sequence 10, Appl
210	5	1.2	17	1	US-08-105-483-158	Sequence 158, App
211	5	1.2	17	1	US-08-709-209-158	Sequence 158, App
212	5	1.2	17	1	US-08-303-275-46	Sequence 46, Appl
213	5	1.2	17	1	US-08-458-101-158	Sequence 158, App
214	5	1.2	17	1	US-08-408-604A-27	Sequence 27, Appl
215	5	1.2	17	4	US-08-896-933-2	Sequence 2, Appli
216	5	1.2	17	4	US-08-896-933-10	Sequence 10, Appl
217	5	1.2	17	4	US-08-896-933-14	Sequence 14, Appl
218	5	1.2	17	4	US-08-896-933-15	Sequence 15, Appl
219	5	1.2	17	4	US-08-905-223-18	Sequence 18, Appl
220	5	1.2	17	5	PCT-US93-09626-27	Sequence 27, Appl
221	5	1.2	19	1	US-07-663-413-8	Sequence 8, Appli
222	5	1.2	19	1	US-07-663-413-11	Sequence 11, Appl
223	5	1.2	19	1	US-07-663-413-30	Sequence 30, Appl
224	5	1.2	19	1	US-08-055-530-8	Sequence 8, Appli
225	5	1.2	19	1	US-08-055-530-11	Sequence 11, Appl
226	5	1.2	19	1	US-08-055-530-30	Sequence 30, Appl
227	5	1.2	19	1	US-08-191-866D-77	Sequence 77, Appl
228	5	1.2	19	2	US-08-185-949B-77	Sequence 77, Appl
229	5	1.2	19	2	US-09-174-060-20	Sequence 20, Appl
230	5	1.2	19	3	US-08-338-382-20	Sequence 20, Appl
231	5	1.2	19	5	PCT-US93-07261-18	Sequence 18, Appl
232	5	1.2	20	1	US-07-956-700B-102	Sequence 102, App
233	5	1.2	20	1	US-08-218-025A-79	Sequence 79, Appl
234	5	1.2	20	1	US-07-678-974D-22	Sequence 22, Appl
235	5	1.2	20	1	US-08-306-116A-31	Sequence 31, Appl
236	5	1.2	20	1	US-08-476-537-102	Sequence 102, App
237	5	1.2	20	1	US-08-484-135-80	Sequence 80, Appl
238	5	1.2	20	1	US-08-484-635-43	Sequence 43, Appl
239	5	1.2	20	1	US-08-485-607-102	Sequence 102, App
240	5	1.2	20	2	US-08-484-631-43	Sequence 43, Appl
241	5	1.2	20	2	US-08-934-915-38	Sequence 38, Appl
242	5	1.2	20	2	US-08-934-915-146	Sequence 146, App
243	5	1.2	20	2	US-08-475-879-102	Sequence 102, App
244	5	1.2	20	2	US-08-827-570-43	Sequence 43, Appl
245	5	1.2	20	2	US-08-945-168-27	Sequence 27, Appl
246	5	1.2	20	3	US-08-935-263-18	Sequence 18, Appl

247	5	1.2	20	5	PCT-US92-07218-2	Sequence 2, Appli
248	5	1.2	21	1	US-08-447-411-58	Sequence 58, Appl
249	5	1.2	21	2	US-08-662-227-15	Sequence 15, Appl
250	5	1.2	21	3	US-08-460-576-16	Sequence 16, Appl
251	5	1.2	21	3	US-08-460-576-17	Sequence 17, Appl
252	5	1.2	21	6	5189145-1	Patent No. 5189145
253	5	1.2	22	1	US-08-080-073-35	Sequence 35, Appl
254	5	1.2	22	3	US-08-658-136-55	Sequence 55, Appl
255	5	1.2	23	1	US-08-248-505-3	Sequence 3, Appli
256	5	1.2	23	4	US-09-587-066-7	Sequence 7, Appli
257	5	1.2	24	1	US-08-103-445-18	Sequence 18, Appl
258	5	1.2	24	1	US-08-103-445-20	Sequence 20, Appl
259	5	1.2	24	1	US-08-461-690B-18	Sequence 18, Appl
260	5	1.2	24	1	US-08-461-690B-20	Sequence 20, Appl
261	5	1.2	24	3	US-08-838-413A-18	Sequence 18, Appl
262	5	1.2	24	3	US-08-838-413A-22	Sequence 22, Appl
263	5	1.2	26	1	US-08-254-493-7	Sequence 7, Appli
264	5	1.2	26	1	US-08-451-472-16	Sequence 16, Appl
265	5	1.2	26	1	US-08-408-222B-7	Sequence 7, Appli
266	5	1.2	26	2	US-08-332-562A-25	Sequence 25, Appl
267	5	1.2	27	2	US-08-637-759B-356	Sequence 356, App
268	5	1.2	27	2	US-08-031-538-58	Sequence 58, Appl
269	5	1.2	27	3	US-08-871-355A-356	Sequence 356, App
270	5	1.2	28	1	US-08-220-378-7	Sequence 7, Appli
271	5	1.2	28	2	US-08-696-012-7	Sequence 7, Appli
272	5	1.2	29	2	US-08-846-762-45	Sequence 45, Appl
273	5	1.2	30	1	US-07-596-081A-35	Sequence 35, Appl
274	5	1.2	30	1	US-07-901-874B-11	Sequence 11, Appl
275	5	1.2	30	1	US-08-457-865-11	Sequence 11, Appl
276	5	1.2	30	3	US-09-045-632-79	Sequence 79, Appl
277	5	1.2	30	3	US-08-486-099-165	Sequence 165, App
278	5	1.2	30	3	US-08-484-223B-165	Sequence 165, App
279	5	1.2	30	3	US-08-919-597-165	Sequence 165, App
280	5	1.2	30	3	US-08-475-668A-165	Sequence 165, App
281	5	1.2	30	3	US-08-485-551A-165	Sequence 165, App
282	5	1.2	30	3	US-08-471-913A-165	Sequence 165, App
283	5	1.2	30	4	US-08-485-264A-165	Sequence 165, App
284	5	1.2	31	1	US-08-097-938-21	Sequence 21, Appl
285	5	1.2	31	1	US-08-476-000-21	Sequence 21, Appl
286	5	1.2	31	1	US-08-472-840-21	Sequence 21, Appl
287	5	1.2	31	2	US-08-476-976-21	Sequence 21, Appl
288	5	1.2	31	3	US-08-519-385B-1	Sequence 1, Appli
289	5	1.2	31	3	US-08-474-410-21	Sequence 21, Appl
290	5	1.2	33	2	US-08-690-011A-38	Sequence 38, Appl
291	5	1.2	34	1	US-07-596-081A-15	Sequence 15, Appl
292	5	1.2	34	2	US-08-690-011A-53	Sequence 53, Appl
293	5	1.2	34	2	US-08-690-011A-54	Sequence 54, Appl
294	5	1.2	34	2	US-08-690-011A-57	Sequence 57, Appl
295	5	1.2	34	3	US-08-746-411A-12	Sequence 12, Appl
296	5	1.2	35	3	US-08-658-136-50	Sequence 50, Appl
297	5	1.2	35	3	US-08-658-136-51	Sequence 51, Appl
298	5	1.2	35	3	US-08-658-136-52	Sequence 52, Appl
299	5	1.2	35	4	US-08-810-009-30	Sequence 30, Appl
300	5	1.2	35	4	US-08-810-009-58	Sequence 58, Appl
301	5	1.2	38	1	US-08-073-028-10	Sequence 10, Appl
302	5	1.2	38	1	US-08-176-500-24	Sequence 24, Appl
303	5	1.2	38	1	US-08-176-500-76	Sequence 76, Appl
304	5	1.2	38	1	US-08-471-052A-24	Sequence 24, Appl
305	5	1.2	38	1	US-08-471-052A-76	Sequence 76, Appl
306	5	1.2	38	1	US-08-374-666-16	Sequence 16, Appl
307	5	1.2	38	1	US-08-189-331-24	Sequence 24, Appl
308	5	1.2	38	1	US-08-189-331-76	Sequence 76, Appl
309	5	1.2	38	2	US-08-471-939-24	Sequence 24, Appl
310	5	1.2	38	2	US-08-471-939-76	Sequence 76, Appl
311	5	1.2	38	2	US-08-471-800-24	Sequence 24, Appl
312	5	1.2	38	2	US-08-471-800-76	Sequence 76, Appl
313	5	1.2	38	2	US-08-488-161-13	Sequence 13, Appl
314	5	1.2	38	2	US-08-471-068-24	Sequence 24, Appl
315	5	1.2	38	2	US-08-471-068-76	Sequence 76, Appl
316	5	1.2	38	3	US-08-486-099-10	Sequence 10, Appl
317	5	1.2	38	3	US-09-273-685-13	Sequence 13, Appl
318	5	1.2	38	3	US-08-360-107A-10	Sequence 10, Appl
319	5	1.2	38	3	US-08-484-223B-10	Sequence 10, Appl
320	5	1.2	38	3	US-08-919-597-10	Sequence 10, Appl
321	5	1.2	38	3	US-08-475-668A-10	Sequence 10, Appl
322	5	1.2	38	3	US-08-485-551A-10	Sequence 10, Appl
323	5	1.2	38	3	US-08-471-913A-10	Sequence 10, Appl
324	5	1.2	38	4	US-08-554-616-10	Sequence 10, Appl
325	5	1.2	38	4	US-08-485-264A-10	Sequence 10, Appl
326	5	1.2	38	5	PCT-US95-11934-13	Sequence 13, Appl
327	5	1.2	40	1	US-07-901-874B-5	Sequence 5, Appli
328	5	1.2	40	1	US-08-457-865-5	Sequence 5, Appli
329	5	1.2	40	2	US-08-557-309B-27	Sequence 27, Appl
330	5	1.2	40	3	US-08-834-306-27	Sequence 27, Appl
331	5	1.2	40	4	US-08-993-674A-27	Sequence 27, Appl
332	5	1.2	41	2	US-08-896-605A-1	Sequence 1, Appli
333	5	1.2	41	2	US-08-896-501A-1	Sequence 1, Appli
334	5	1.2	41	4	US-08-706-344C-23	Sequence 23, Appl
335	5	1.2	42	1	US-08-444-005-24	Sequence 24, Appl
336	5	1.2	42	1	US-08-444-005-25	Sequence 25, Appl
337	5	1.2	42	4	US-08-974-549A-141	Sequence 141, App
338	5	1.2	43	1	US-08-178-477B-21	Sequence 21, Appl
339	5	1.2	45	2	US-08-766-858A-20	Sequence 20, Appl
340	5	1.2	45	4	US-08-796-899-32	Sequence 32, Appl
341	5	1.2	47	1	US-08-244-491A-3	Sequence 3, Appli
342	5	1.2	47	2	US-08-797-727A-3	Sequence 3, Appli
343	5	1.2	48	2	US-08-469-537A-60	Sequence 60, Appl
344	5	1.2	48	2	US-08-469-537A-68	Sequence 68, Appl
345	5	1.2	48	4	US-08-905-223-486	Sequence 486, App
346	5	1.2	52	2	US-08-245-511-44	Sequence 44, Appl
347	5	1.2	52	2	US-08-600-993A-44	Sequence 44, Appl
348	5	1.2	52	4	US-09-082-593-8	Sequence 8, Appli
349	5	1.2	53	2	US-08-676-782-9	Sequence 9, Appli
350	5	1.2	53	3	US-08-461-607-15	Sequence 15, Appl
351	5	1.2	53	4	US-08-905-223-326	Sequence 326, App
352	5	1.2	53	4	US-09-363-600-15	Sequence 15, Appl
353	5	1.2	54	1	US-08-159-340A-39	Sequence 39, Appl
354	5	1.2	54	2	US-08-660-963-16	Sequence 16, Appl
355	5	1.2	54	2	US-08-660-963-17	Sequence 17, Appl
356	5	1.2	54	2	US-08-660-963-18	Sequence 18, Appl
357	5	1.2	56	6	5466585-2	Patent No. 5466585
358	5	1.2	57	5	PCT-US94-05150-22	Sequence 22, Appl
359	5	1.2	57	6	5466585-3	Patent No. 5466585
360	5	1.2	59	3	US-09-065-474-168	Sequence 168, App
361	5	1.2	59	4	US-09-300-672-4	Sequence 4, Appli
362	5	1.2	60	4	US-08-928-213B-72	Sequence 72, Appl
363	5	1.2	61	4	US-09-416-050A-30	Sequence 30, Appl
364	5	1.2	61	4	US-09-416-050A-31	Sequence 31, Appl
365	5	1.2	61	4	US-09-664-800-30	Sequence 30, Appl
366	5	1.2	61	4	US-09-664-800-31	Sequence 31, Appl
367	5	1.2	61	4	US-09-665-309-30	Sequence 30, Appl
368	5	1.2	61	4	US-09-665-309-31	Sequence 31, Appl
369	5	1.2	61	4	US-09-661-569-30	Sequence 30, Appl
370	5	1.2	61	4	US-09-661-569-31	Sequence 31, Appl
371	5	1.2	62	2	US-08-385-590A-16	Sequence 16, Appl
372	5	1.2	62	3	US-09-021-520-16	Sequence 16, Appl
373	5	1.2	63	2	US-08-385-590A-14	Sequence 14, Appl
374	5	1.2	63	3	US-09-021-520-14	Sequence 14, Appl
375	5	1.2	64	2	US-08-209-521-20	Sequence 20, Appl
376	5	1.2	64	4	US-08-961-810-130	Sequence 130, App
377	5	1.2	64	4	US-08-352-902D-130	Sequence 130, App
378	5	1.2	66	4	US-09-082-593-6	Sequence 6, Appli
379	5	1.2	66	4	US-09-007-905-41	Sequence 41, Appl
380	5	1.2	69	4	US-08-905-223-419	Sequence 419, App
381	5	1.2	70	1	US-07-956-700B-25	Sequence 25, Appl
382	5	1.2	70	1	US-08-476-537-25	Sequence 25, Appl
383	5	1.2	70	1	US-08-485-607-25	Sequence 25, Appl
384	5	1.2	70	2	US-08-475-879-25	Sequence 25, Appl
385	5	1.2	75	1	US-07-934-017-1	Sequence 1, Appli
386	5	1.2	75	3	US-09-042-012-8	Sequence 8, Appli
387	5	1.2	75	4	US-09-305-086-2	Sequence 2, Appli
388	5	1.2	76	1	US-08-168-941-1	Sequence 1, Appli
389	5	1.2	76	2	US-08-117-952-753	Sequence 753, App
390	5	1.2	77	2	US-08-436-420-44	Sequence 44, Appl
391	5	1.2	77	4	US-08-995-159-7	Sequence 7, Appli
392	5	1.2	78	2	US-08-465-380-28	Sequence 28, Appl

393	5	1.2	78	2	US-08-465-380-53	Sequence 53, Appl
394	5	1.2	78	2	US-08-465-380-54	Sequence 54, Appl
395	5	1.2	78	2	US-08-465-380-56	Sequence 56, Appl
396	5	1.2	78	2	US-08-480-478-57	Sequence 57, Appl
397	5	1.2	78	2	US-08-486-397-28	Sequence 28, Appl
398	5	1.2	78	2	US-08-486-397-53	Sequence 53, Appl
399	5	1.2	78	2	US-08-486-397-54	Sequence 54, Appl
400	5	1.2	78	2	US-08-486-397-56	Sequence 56, Appl
401	5	1.2	78	2	US-08-486-399-28	Sequence 28, Appl
402	5	1.2	78	2	US-08-486-399-53	Sequence 53, Appl
403	5	1.2	78	2	US-08-486-399-54	Sequence 54, Appl
404	5	1.2	78	2	US-08-486-399-56	Sequence 56, Appl
405	5	1.2	78	2	US-08-461-965-28	Sequence 28, Appl
406	5	1.2	78	2	US-08-461-965-53	Sequence 53, Appl
407	5	1.2	78	2	US-08-461-965-54	Sequence 54, Appl
408	5	1.2	78	2	US-08-461-965-56	Sequence 56, Appl
409	5	1.2	78	2	US-08-326-110A-57	Sequence 57, Appl
410	5	1.2	78	2	US-08-634-641-28	Sequence 28, Appl
411	5	1.2	78	2	US-08-634-641-53	Sequence 53, Appl
412	5	1.2	78	2	US-08-634-641-54	Sequence 54, Appl
413	5	1.2	78	2	US-08-634-641-56	Sequence 56, Appl
414	5	1.2	78	3	US-09-249-471-28	Sequence 28, Appl
415	5	1.2	78	3	US-09-249-471-53	Sequence 53, Appl
416	5	1.2	78	3	US-09-249-471-54	Sequence 54, Appl
417	5	1.2	78	3	US-09-249-471-56	Sequence 56, Appl
418	5	1.2	78	3	US-09-249-472-28	Sequence 28, Appl
419	5	1.2	78	3	US-09-249-472-53	Sequence 53, Appl
420	5	1.2	78	3	US-09-249-472-54	Sequence 54, Appl
421	5	1.2	78	3	US-09-249-472-56	Sequence 56, Appl
422	5	1.2	78	3	US-09-249-451-28	Sequence 28, Appl
423	5	1.2	78	3	US-09-249-451-53	Sequence 53, Appl
424	5	1.2	78	3	US-09-249-451-54	Sequence 54, Appl
425	5	1.2	78	3	US-09-249-451-56	Sequence 56, Appl
426	5	1.2	78	3	US-08-809-455-28	Sequence 28, Appl
427	5	1.2	78	3	US-08-809-455-53	Sequence 53, Appl
428	5	1.2	78	3	US-08-809-455-54	Sequence 54, Appl
429	5	1.2	78	3	US-08-809-455-56	Sequence 56, Appl
430	5	1.2	78	3	US-09-249-461-28	Sequence 28, Appl
431	5	1.2	78	3	US-09-249-461-53	Sequence 53, Appl
432	5	1.2	78	3	US-09-249-461-54	Sequence 54, Appl
433	5	1.2	78	3	US-09-249-461-56	Sequence 56, Appl
434	5	1.2	78	3	US-09-249-448-28	Sequence 28, Appl
435	5	1.2	78	3	US-09-249-448-53	Sequence 53, Appl
436	5	1.2	78	3	US-09-249-448-54	Sequence 54, Appl
437	5	1.2	78	3	US-09-249-448-56	Sequence 56, Appl
438	5	1.2	79	4	US-09-025-151-10	Sequence 10, Appl
439	5	1.2	80	1	US-08-100-247-3	Sequence 3, Appli
440	5	1.2	80	1	US-08-483-146A-3	Sequence 3, Appli
441	5	1.2	80	1	US-08-232-513A-4	Sequence 4, Appli
442	5	1.2	80	1	US-08-484-594A-3	Sequence 3, Appli
443	5	1.2	80	2	US-08-637-759B-472	Sequence 472, App
444	5	1.2	80	3	US-08-871-355A-472	Sequence 472, App
445	5	1.2	82	1	US-07-881-075-19	Sequence 19, Appl
446	5	1.2	82	1	US-08-120-827-19	Sequence 19, Appl
447	5	1.2	82	1	US-08-478-675-19	Sequence 19, Appl
448	5	1.2	82	4	US-09-188-930-301	Sequence 301, App
449	5	1.2	82	4	US-09-091-590A-3	Sequence 3, Appli
450	5	1.2	83	1	US-07-881-075-17	Sequence 17, Appl
451	5	1.2	83	1	US-08-120-827-17	Sequence 17, Appl
452	5	1.2	83	1	US-08-478-675-17	Sequence 17, Appl
453	5	1.2	87	1	US-07-881-075-18	Sequence 18, Appl
454	5	1.2	87	1	US-08-120-827-18	Sequence 18, Appl
455	5	1.2	87	1	US-08-478-675-18	Sequence 18, Appl
456	5	1.2	87	4	US-09-042-785A-29	Sequence 29, Appl
457	5	1.2	88	2	US-08-690-011A-11	Sequence 11, Appl
458	5	1.2	88	3	US-09-045-632-7	Sequence 7, Appli
459	5	1.2	89	2	US-08-690-011A-52	Sequence 52, Appl
460	5	1.2	90	1	US-08-370-225-26	Sequence 26, Appl
461	5	1.2	90	1	US-08-464-339A-12	Sequence 12, Appl
462	5	1.2	90	1	US-08-461-859-26	Sequence 26, Appl
463	5	1.2	90	2	US-08-630-822A-58	Sequence 58, Appl
464	5	1.2	90	2	US-08-630-822A-80	Sequence 80, Appl
465	5	1.2	90	2	US-09-005-069-58	Sequence 58, Appl
466	5	1.2	90	2	US-09-005-069-80	Sequence 80, Appl
467	5	1.2	90	5	PCT-US93-10069-26	Sequence 26, Appl
468	5	1.2	91	1	US-08-451-947-49	Sequence 49, Appl
469	5	1.2	91	2	US-08-424-826A-49	Sequence 49, Appl
470	5	1.2	91	3	US-08-928-694-49	Sequence 49, Appl
471	5	1.2	91	5	PCT-US91-06950-49	Sequence 49, Appl
472	5	1.2	92	1	US-07-792-988-2	Sequence 2, Appli
473	5	1.2	92	2	US-08-485-937-13	Sequence 13, Appl
474	5	1.2	92	2	US-08-373-215-13	Sequence 13, Appl
475	5	1.2	92	5	PCT-US93-06552-13	Sequence 13, Appl
476	5	1.2	95	3	US-09-024-023-4	Sequence 4, Appli
477	5	1.2	97	1	US-08-202-389-16	Sequence 16, Appl
478	5	1.2	98	1	US-07-795-859B-8	Sequence 8, Appli
479	5	1.2	98	1	US-08-457-616-8	Sequence 8, Appli
480	5	1.2	98	2	US-08-690-011A-19	Sequence 19, Appl
481	5	1.2	98	2	US-09-047-125-15	Sequence 15, Appl
482	5	1.2	98	3	US-07-736-335E-15	Sequence 15, Appl
483	5	1.2	98	4	US-09-326-135-4	Sequence 4, Appli
484	5	1.2	99	1	US-08-036-210-26	Sequence 26, Appl
485	5	1.2	99	2	US-08-449-609-26	Sequence 26, Appl
486	5	1.2	100	1	US-08-307-499-3	Sequence 3, Appli
487	5	1.2	100	4	US-09-299-268-3	Sequence 3, Appli
488	5	1.2	101	4	US-08-905-223-443	Sequence 443, App
489	5	1.2	102	4	US-09-232-522-4	Sequence 4, Appli
490	5	1.2	103	1	US-08-271-562-1	Sequence 1, Appli
491	5	1.2	103	1	US-08-087-007-3	Sequence 3, Appli
492	5	1.2	103	2	US-08-696-777-1	Sequence 1, Appli
493	5	1.2	103	3	US-08-483-433-3	Sequence 3, Appli
494	5	1.2	103	5	PCT-US92-05920-3	Sequence 3, Appli
495	5	1.2	104	1	US-07-942-245-16	Sequence 16, Appl
496	5	1.2	104	1	US-08-716-588-1	Sequence 1, Appli
497	5	1.2	104	2	US-08-966-444-1	Sequence 1, Appli
498	5	1.2	106	2	US-08-820-825-14	Sequence 14, Appl
499	5	1.2	106	4	US-09-307-817-14	Sequence 14, Appl
500	5	1.2	107	1	US-07-942-245-14	Sequence 14, Appl
501	5	1.2	107	1	US-08-409-731A-2	Sequence 2, Appli
502	5	1.2	107	2	US-08-470-298B-2	Sequence 2, Appli
503	5	1.2	107	2	US-09-023-073A-2	Sequence 2, Appli
504	5	1.2	107	4	US-09-232-522-2	Sequence 2, Appli
505	5	1.2	108	4	US-09-177-249-172	Sequence 172, App
506	5	1.2	109	3	US-09-053-197A-16	Sequence 16, Appl
507	5	1.2	110	1	US-08-017-570-2	Sequence 2, Appli
508	5	1.2	110	1	US-08-844-064-8	Sequence 8, Appli
509	5	1.2	110	1	US-08-471-426-2	Sequence 2, Appli
510	5	1.2	110	3	US-09-009-433-8	Sequence 8, Appli
511	5	1.2	110	5	PCT-US94-01709-2	Sequence 2, Appli
512	5	1.2	111	3	US-09-040-508-4	Sequence 4, Appli
513	5	1.2	113	1	US-08-077-252B-4	Sequence 4, Appli
514	5	1.2	113	1	US-08-107-669D-2	Sequence 2, Appli
515	5	1.2	113	1	US-08-107-669D-5	Sequence 5, Appli
516	5	1.2	113	1	US-08-472-788A-2	Sequence 2, Appli
517	5	1.2	113	1	US-08-472-788A-5	Sequence 5, Appli
518	5	1.2	113	2	US-08-477-531B-2	Sequence 2, Appli
519	5	1.2	113	2	US-08-477-531B-5	Sequence 5, Appli
520	5	1.2	113	2	US-08-710-330A-5	Sequence 5, Appli
521	5	1.2	113	2	US-08-082-842A-2	Sequence 2, Appli
522	5	1.2	113	2	US-08-082-842A-5	Sequence 5, Appli
523	5	1.2	113	4	US-09-002-753A-4	Sequence 4, Appli
524	5	1.2	113	5	PCT-US94-06687-4	Sequence 4, Appli
525	5	1.2	114	4	US-08-929-856-66	Sequence 66, Appl
526	5	1.2	114	4	US-08-929-856-189	Sequence 189, App
527	5	1.2	115	1	US-07-942-245-7	Sequence 7, Appli
528	5	1.2	117	1	US-07-752-101A-37	Sequence 37, Appl
529	5	1.2	117	1	US-07-752-101A-40	Sequence 40, Appl
530	5	1.2	117	1	US-07-752-101A-42	Sequence 42, Appl
531	5	1.2	117	1	US-08-812-025-7	Sequence 7, Appli
532	5	1.2	118	2	US-08-417-174-2	Sequence 2, Appli
533	5	1.2	118	2	US-08-231-565A-2	Sequence 2, Appli
534	5	1.2	118	2	US-09-007-961-2	Sequence 2, Appli
535	5	1.2	118	3	US-08-545-809A-108	Sequence 108, App
536	5	1.2	119	1	US-08-207-996-20	Sequence 20, Appl
537	5	1.2	119	1	US-08-207-996-27	Sequence 27, Appl
538	5	1.2	119	2	US-08-760-840A-20	Sequence 20, Appl

539	5	1.2	119	2	US-08-760-840A-28	Sequence 28, Appl
540	5	1.2	119	3	US-08-545-809A-131	Sequence 131, App
541	5	1.2	119	3	US-09-266-119-20	Sequence 20, Appl
542	5	1.2	119	3	US-09-266-119-28	Sequence 28, Appl
543	5	1.2	120	1	US-08-347-492B-2	Sequence 2, Appli
544	5	1.2	120	2	US-08-798-143-2	Sequence 2, Appli
545	5	1.2	120	5	PCT-US95-15484-2	Sequence 2, Appli
546	5	1.2	121	1	US-08-111-080-20	Sequence 20, Appl
547	5	1.2	121	1	US-08-275-053-9	Sequence 9, Appli
548	5	1.2	121	1	US-08-450-257-51	Sequence 51, Appl
549	5	1.2	121	1	US-08-450-257-55	Sequence 55, Appl
550	5	1.2	121	1	US-08-211-980-20	Sequence 20, Appl
551	5	1.2	121	1	US-08-450-246-51	Sequence 51, Appl
552	5	1.2	121	1	US-08-450-246-55	Sequence 55, Appl
553	5	1.2	121	1	US-08-450-098-51	Sequence 51, Appl
554	5	1.2	121	1	US-08-450-098-55	Sequence 55, Appl
555	5	1.2	121	1	US-08-451-233-51	Sequence 51, Appl
556	5	1.2	121	1	US-08-451-233-55	Sequence 55, Appl
557	5	1.2	121	1	US-08-450-236-51	Sequence 51, Appl
558	5	1.2	121	1	US-08-450-236-55	Sequence 55, Appl
559	5	1.2	121	5	PCT-US93-07967-20	Sequence 20, Appl
560	5	1.2	122	1	US-07-956-700B-9	Sequence 9, Appli
561	5	1.2	122	1	US-08-476-537-9	Sequence 9, Appli
562	5	1.2	122	1	US-08-485-607-9	Sequence 9, Appli
563	5	1.2	122	2	US-08-475-879-9	Sequence 9, Appli
564	5	1.2	123	1	US-08-094-128A-25	Sequence 25, Appl
565	5	1.2	123	1	US-08-455-674-25	Sequence 25, Appl
566	5	1.2	123	1	US-08-455-992-25	Sequence 25, Appl
567	5	1.2	123	1	US-08-455-972-25	Sequence 25, Appl
568	5	1.2	123	1	US-08-530-010-25	Sequence 25, Appl
569	5	1.2	123	2	US-08-484-101B-25	Sequence 25, Appl
570	5	1.2	123	5	PCT-US92-00652-25	Sequence 25, Appl
571	5	1.2	124	4	US-09-082-593-4	Sequence 4, Appli
572	5	1.2	124	4	US-09-240-274-2	Sequence 2, Appli
573	5	1.2	124	4	US-09-240-274-3	Sequence 3, Appli
574	5	1.2	124	4	US-09-240-274-4	Sequence 4, Appli
575	5	1.2	124	4	US-09-240-274-7	Sequence 7, Appli
576	5	1.2	124	6	5466585-7	Patent No. 5466585
577	5	1.2	125	1	US-08-313-075A-54	Sequence 54, Appl
578	5	1.2	125	2	US-08-164-292B-8	Sequence 8, Appli
579	5	1.2	125	3	US-08-845-623-8	Sequence 8, Appli
580	5	1.2	125	3	US-08-815-927-8	Sequence 8, Appli
581	5	1.2	126	1	US-08-312-387B-2	Sequence 2, Appli
582	5	1.2	126	1	US-08-683-426-2	Sequence 2, Appli
583	5	1.2	126	1	US-08-683-458-2	Sequence 2, Appli
584	5	1.2	126	2	US-08-878-360-2	Sequence 2, Appli
585	5	1.2	126	3	US-08-478-140B-2	Sequence 2, Appli
586	5	1.2	128	6	5179198-1	Patent No. 5179198
587	5	1.2	128	6	5521296-1	Patent No. 5521296
588	5	1.2	130	1	US-07-648-796A-4	Sequence 4, Appli
589	5	1.2	130	1	US-08-686-878A-9	Sequence 9, Appli
590	5	1.2	132	1	US-07-820-154A-2	Sequence 2, Appli
591	5	1.2	132	1	US-07-820-154A-6	Sequence 6, Appli
592	5	1.2	132	2	US-08-097-554A-2	Sequence 2, Appli
593	5	1.2	132	2	US-08-097-554A-6	Sequence 6, Appli
594	5	1.2	132	3	US-08-480-640A-2	Sequence 2, Appli
595	5	1.2	132	3	US-08-480-640A-6	Sequence 6, Appli
596	5	1.2	132	3	US-08-295-802-2	Sequence 2, Appli
597	5	1.2	132	3	US-08-295-802-6	Sequence 6, Appli
598	5	1.2	132	4	US-08-686-968C-102	Sequence 102, App
599	5	1.2	132	4	US-08-488-237A-2	Sequence 2, Appli
600	5	1.2	132	4	US-08-488-237A-6	Sequence 6, Appli
601	5	1.2	132	5	PCT-US93-00324-2	Sequence 2, Appli
602	5	1.2	132	5	PCT-US93-00324-6	Sequence 6, Appli
603	5	1.2	134	1	US-08-111-080-16	Sequence 16, Appl
604	5	1.2	134	1	US-08-450-257-38	Sequence 38, Appl
605	5	1.2	134	1	US-08-211-980-16	Sequence 16, Appl
606	5	1.2	134	1	US-08-450-246-38	Sequence 38, Appl
607	5	1.2	134	1	US-08-450-098-38	Sequence 38, Appl
608	5	1.2	134	1	US-08-451-233-38	Sequence 38, Appl
609	5	1.2	134	1	US-08-450-236-38	Sequence 38, Appl
610	5	1.2	134	3	US-09-094-287-2	Sequence 2, Appli
611	5	1.2	134	4	US-09-251-545-2	Sequence 2, Appli
612	5	1.2	134	5	PCT-US92-07111-15	Sequence 15, Appl
613	5	1.2	134	5	PCT-US93-07967-16	Sequence 16, Appl
614	5	1.2	135	2	US-08-820-825-2	Sequence 2, Appli
615	5	1.2	135	3	US-08-899-031-1	Sequence 1, Appli
616	5	1.2	135	4	US-08-985-950-10	Sequence 10, Appl
617	5	1.2	135	4	US-09-307-817-2	Sequence 2, Appli
618	5	1.2	136	5	PCT-US95-07171-2	Sequence 2, Appli
619	5	1.2	137	2	US-08-933-750C-49	Sequence 49, Appl
620	5	1.2	137	3	US-09-125-265-2	Sequence 2, Appli
621	5	1.2	137	4	US-09-234-613-49	Sequence 49, Appl
622	5	1.2	137	4	US-09-254-722-1	Sequence 1, Appli
623	5	1.2	137	4	US-09-381-737-1	Sequence 1, Appli
624	5	1.2	138	3	US-09-173-581-9	Sequence 9, Appli
625	5	1.2	139	1	US-07-648-796A-3	Sequence 3, Appli
626	5	1.2	140	2	US-08-449-287-4	Sequence 4, Appli
627	5	1.2	141	2	US-08-765-783A-29	Sequence 29, Appl
628	5	1.2	141	3	US-08-921-100-29	Sequence 29, Appl
629	5	1.2	141	3	US-08-880-142-29	Sequence 29, Appl
630	5	1.2	141	3	US-08-902-201-29	Sequence 29, Appl
631	5	1.2	141	4	US-09-416-557-29	Sequence 29, Appl
632	5	1.2	142	4	US-09-269-258A-4	Sequence 4, Appli
633	5	1.2	143	1	US-08-450-257-63	Sequence 63, Appl
634	5	1.2	143	1	US-08-450-246-63	Sequence 63, Appl
635	5	1.2	143	1	US-08-450-098-63	Sequence 63, Appl
636	5	1.2	143	1	US-08-451-233-63	Sequence 63, Appl
637	5	1.2	143	1	US-08-450-236-63	Sequence 63, Appl
638	5	1.2	144	1	US-08-026-320A-2	Sequence 2, Appli
639	5	1.2	145	2	US-08-540-242A-6	Sequence 6, Appli
640	5	1.2	145	2	US-08-394-021-11	Sequence 11, Appl
641	5	1.2	145	3	US-09-094-287-4	Sequence 4, Appli
642	5	1.2	145	5	PCT-US96-01471-6	Sequence 6, Appli
643	5	1.2	146	1	US-07-795-859B-26	Sequence 26, Appl
644	5	1.2	146	1	US-08-240-712-24	Sequence 24, Appl
645	5	1.2	146	1	US-08-457-616-26	Sequence 26, Appl
646	5	1.2	146	1	US-08-443-890-24	Sequence 24, Appl
647	5	1.2	146	2	US-08-484-686B-66	Sequence 66, Appl
648	5	1.2	146	2	US-09-003-081-3	Sequence 3, Appli
649	5	1.2	146	3	US-08-648-262-3	Sequence 3, Appli
650	5	1.2	146	3	US-08-648-263-3	Sequence 3, Appli
651	5	1.2	146	4	US-08-463-160B-67	Sequence 67, Appl
652	5	1.2	146	4	US-09-058-562-22	Sequence 22, Appl
653	5	1.2	146	5	PCT-US92-09752-24	Sequence 24, Appl
654	5	1.2	147	3	US-08-485-942A-95	Sequence 95, Appl
655	5	1.2	147	3	US-08-488-214A-95	Sequence 95, Appl
656	5	1.2	147	3	US-08-488-208A-95	Sequence 95, Appl
657	5	1.2	148	1	US-08-450-257-59	Sequence 59, Appl
658	5	1.2	148	1	US-08-450-246-59	Sequence 59, Appl
659	5	1.2	148	1	US-08-450-098-59	Sequence 59, Appl
660	5	1.2	148	1	US-08-451-233-59	Sequence 59, Appl
661	5	1.2	148	1	US-08-468-347-22	Sequence 22, Appl
662	5	1.2	148	1	US-08-450-236-59	Sequence 59, Appl
663	5	1.2	148	2	US-08-467-389-22	Sequence 22, Appl
664	5	1.2	148	2	US-08-779-379-22	Sequence 22, Appl
665	5	1.2	148	2	US-08-469-219-22	Sequence 22, Appl
666	5	1.2	148	4	US-09-228-152-22	Sequence 22, Appl
667	5	1.2	149	2	US-08-039-364-18	Sequence 18, Appl
668	5	1.2	149	4	US-08-975-762-63	Sequence 63, Appl
669	5	1.2	149	4	US-09-158-710-18	Sequence 18, Appl
670	5	1.2	151	1	US-08-253-155A-44	Sequence 44, Appl
671	5	1.2	153	1	US-08-226-264-26	Sequence 26, Appl
672	5	1.2	153	3	US-08-840-146-10	Sequence 10, Appl
673	5	1.2	153	3	US-09-360-220-10	Sequence 10, Appl
674	5	1.2	154	2	US-08-330-394A-29	Sequence 29, Appl
675	5	1.2	155	1	US-07-940-605A-6	Sequence 6, Appli
676	5	1.2	155	2	US-08-690-096-6	Sequence 6, Appli
677	5	1.2	156	1	US-08-844-064-10	Sequence 10, Appl
678	5	1.2	156	1	US-08-074-121-5	Sequence 5, Appli
679	5	1.2	156	2	US-08-330-394A-22	Sequence 22, Appl
680	5	1.2	156	3	US-09-009-433-10	Sequence 10, Appl
681	5	1.2	156	4	US-09-181-183-2	Sequence 2, Appli
682	5	1.2	156	5	PCT-US94-06447-5	Sequence 5, Appli
683	5	1.2	157	1	US-08-257-999-4	Sequence 4, Appli
684	5	1.2	157	1	US-08-450-257-60	Sequence 60, Appl

685	5	1.2	157	1	US-08-450-246-60	Sequence 60, Appl
686	5	1.2	157	1	US-08-450-098-60	Sequence 60, Appl
687	5	1.2	157	1	US-08-451-233-60	Sequence 60, Appl
688	5	1.2	157	1	US-08-494-577-4	Sequence 4, Appli
689	5	1.2	157	1	US-08-494-577-5	Sequence 5, Appli
690	5	1.2	157	1	US-08-450-065-2	Sequence 2, Appli
691	5	1.2	157	1	US-08-450-595-2	Sequence 2, Appli
692	5	1.2	157	1	US-08-450-236-60	Sequence 60, Appl
693	5	1.2	157	2	US-08-795-868-4	Sequence 4, Appli
694	5	1.2	157	2	US-08-795-868-5	Sequence 5, Appli
695	5	1.2	157	4	US-09-181-183-4	Sequence 4, Appli
696	5	1.2	158	2	US-08-670-186-4	Sequence 4, Appli
697	5	1.2	158	2	US-08-933-750C-15	Sequence 15, Appl
698	5	1.2	158	4	US-09-234-613-15	Sequence 15, Appl
699	5	1.2	159	1	US-08-253-155A-27	Sequence 27, Appl
700	5	1.2	159	1	US-08-597-236-9	Sequence 9, Appli
701	5	1.2	159	1	US-08-746-682A-9	Sequence 9, Appli
702	5	1.2	159	2	US-08-938-717-2	Sequence 2, Appli
703	5	1.2	159	2	US-08-810-655A-10	Sequence 10, Appl
704	5	1.2	159	3	US-09-310-488-2	Sequence 2, Appli
705	5	1.2	160	4	US-09-433-428D-70	Sequence 70, Appl
706	5	1.2	160	4	US-09-188-930-321	Sequence 321, App
707	5	1.2	160	4	US-09-587-066-4	Sequence 4, Appli
708	5	1.2	161	2	US-08-465-380-65	Sequence 65, Appl
709	5	1.2	161	2	US-08-486-397-65	Sequence 65, Appl
710	5	1.2	161	2	US-08-486-399-65	Sequence 65, Appl
711	5	1.2	161	2	US-08-461-965-65	Sequence 65, Appl
712	5	1.2	161	2	US-08-403-852D-30	Sequence 30, Appl
713	5	1.2	161	2	US-08-634-641-65	Sequence 65, Appl
714	5	1.2	161	3	US-09-249-471-65	Sequence 65, Appl
715	5	1.2	161	3	US-09-249-472-65	Sequence 65, Appl
716	5	1.2	161	3	US-08-510-646B-31	Sequence 31, Appl
717	5	1.2	161	3	US-09-249-451-65	Sequence 65, Appl
718	5	1.2	161	3	US-08-809-455-65	Sequence 65, Appl
719	5	1.2	161	3	US-09-249-461-65	Sequence 65, Appl
720	5	1.2	161	3	US-09-249-448-65	Sequence 65, Appl
721	5	1.2	161	4	US-09-231-818-30	Sequence 30, Appl
722	5	1.2	162	1	US-08-048-164A-2	Sequence 2, Appli
723	5	1.2	162	1	US-08-460-462-2	Sequence 2, Appli
724	5	1.2	162	1	US-08-460-457-2	Sequence 2, Appli
725	5	1.2	162	1	US-08-460-458-2	Sequence 2, Appli
726	5	1.2	162	2	US-08-460-455-2	Sequence 2, Appli
727	5	1.2	162	2	US-08-330-394A-2	Sequence 2, Appli
728	5	1.2	162	2	US-08-465-380-63	Sequence 63, Appl
729	5	1.2	162	2	US-08-465-380-64	Sequence 64, Appl
730	5	1.2	162	2	US-08-486-397-63	Sequence 63, Appl
731	5	1.2	162	2	US-08-486-397-64	Sequence 64, Appl
732	5	1.2	162	2	US-08-486-399-63	Sequence 63, Appl
733	5	1.2	162	2	US-08-486-399-64	Sequence 64, Appl
734	5	1.2	162	2	US-08-461-965-63	Sequence 63, Appl
735	5	1.2	162	2	US-08-461-965-64	Sequence 64, Appl
736	5	1.2	162	2	US-08-634-641-63	Sequence 63, Appl
737	5	1.2	162	2	US-08-634-641-64	Sequence 64, Appl
738	5	1.2	162	3	US-09-249-471-63	Sequence 63, Appl
739	5	1.2	162	3	US-09-249-471-64	Sequence 64, Appl
740	5	1.2	162	3	US-09-249-472-63	Sequence 63, Appl
741	5	1.2	162	3	US-09-249-472-64	Sequence 64, Appl
742	5	1.2	162	3	US-09-249-451-63	Sequence 63, Appl
743	5	1.2	162	3	US-09-249-451-64	Sequence 64, Appl
744	5	1.2	162	3	US-08-809-455-63	Sequence 63, Appl
745	5	1.2	162	3	US-08-809-455-64	Sequence 64, Appl
746	5	1.2	162	3	US-09-249-461-63	Sequence 63, Appl
747	5	1.2	162	3	US-09-249-461-64	Sequence 64, Appl
748	5	1.2	162	3	US-09-249-448-63	Sequence 63, Appl
749	5	1.2	162	3	US-09-249-448-64	Sequence 64, Appl
750	5	1.2	163	1	US-08-075-545-1	Sequence 1, Appli
751	5	1.2	163	1	US-08-256-261-17	Sequence 17, Appl
752	5	1.2	163	3	US-09-006-636-7	Sequence 7, Appli
753	5	1.2	163	3	US-08-852-299-17	Sequence 17, Appl
754	5	1.2	163	4	US-08-961-083-50	Sequence 50, Appl
755	5	1.2	163	4	US-09-006-632-7	Sequence 7, Appli
756	5	1.2	163	5	PCT-US95-08295-10	Sequence 10, Appl
757	5	1.2	164	2	US-08-918-723-3	Sequence 3, Appli
758	5	1.2	164	2	US-09-237-507-3	Sequence 3, Appli
759	5	1.2	166	2	US-08-347-563A-5	Sequence 5, Appli
760	5	1.2	166	3	US-08-292-345B-5	Sequence 5, Appli
761	5	1.2	166	3	US-08-485-942A-5	Sequence 5, Appli
762	5	1.2	166	3	US-08-488-214A-5	Sequence 5, Appli
763	5	1.2	166	3	US-08-488-208A-5	Sequence 5, Appli
764	5	1.2	167	1	US-07-918-953-8	Sequence 8, Appli
765	5	1.2	167	1	US-08-081-661-8	Sequence 8, Appli
766	5	1.2	167	1	US-07-688-352C-26	Sequence 26, Appl
767	5	1.2	167	2	US-08-540-242A-2	Sequence 2, Appli
768	5	1.2	167	2	US-08-347-563A-2	Sequence 2, Appli
769	5	1.2	167	2	US-08-690-849-2	Sequence 2, Appli
770	5	1.2	167	2	US-09-003-081-2	Sequence 2, Appli
771	5	1.2	167	2	US-08-474-379C-26	Sequence 26, Appl
772	5	1.2	167	3	US-08-292-345B-2	Sequence 2, Appli
773	5	1.2	167	3	US-08-648-262-2	Sequence 2, Appli
774	5	1.2	167	3	US-08-648-263-2	Sequence 2, Appli
775	5	1.2	167	3	US-08-485-942A-2	Sequence 2, Appli
776	5	1.2	167	3	US-09-004-053-2	Sequence 2, Appli
777	5	1.2	167	3	US-09-146-249A-26	Sequence 26, Appl
778	5	1.2	167	3	US-08-206-188B-26	Sequence 26, Appl
779	5	1.2	167	3	US-08-488-214A-2	Sequence 2, Appli
780	5	1.2	167	3	US-08-488-208A-2	Sequence 2, Appli
781	5	1.2	167	4	US-09-062-440-8	Sequence 8, Appli
782	5	1.2	167	4	US-09-062-440-9	Sequence 9, Appli
783	5	1.2	167	4	US-09-062-440-11	Sequence 11, Appl
784	5	1.2	167	4	US-08-759-628-2	Sequence 2, Appli
785	5	1.2	167	5	PCT-US96-01471-2	Sequence 2, Appli
786	5	1.2	168	3	US-09-188-579-85	Sequence 85, Appl
787	5	1.2	168	4	US-09-315-444-85	Sequence 85, Appl
788	5	1.2	169	1	US-08-233-788A-53	Sequence 53, Appl
789	5	1.2	169	2	US-08-766-551-5	Sequence 5, Appli
790	5	1.2	170	1	US-08-127-499A-20	Sequence 20, Appl
791	5	1.2	170	1	US-08-482-847-20	Sequence 20, Appl
792	5	1.2	170	3	US-09-186-250-4	Sequence 4, Appli
793	5	1.2	170	3	US-09-186-250-8	Sequence 8, Appli
794	5	1.2	170	4	US-08-875-573-6	Sequence 6, Appli
795	5	1.2	170	4	US-09-517-347-4	Sequence 4, Appli
796	5	1.2	170	4	US-09-517-347-8	Sequence 8, Appli
797	5	1.2	171	1	US-08-313-681A-4	Sequence 4, Appli
798	5	1.2	171	2	US-08-853-659A-54	Sequence 54, Appl
799	5	1.2	171	3	US-09-322-911-4	Sequence 4, Appli
800	5	1.2	171	4	US-09-188-930-197	Sequence 197, App
801	5	1.2	173	3	US-08-772-270A-10	Sequence 10, Appl
802	5	1.2	174	4	US-09-383-586-12	Sequence 12, Appl
803	5	1.2	175	2	US-08-809-267-5	Sequence 5, Appli
804	5	1.2	175	4	US-08-827-171B-5	Sequence 5, Appli
805	5	1.2	175	5	PCT-US95-13662A-5	Sequence 5, Appli
806	5	1.2	176	4	US-09-188-930-120	Sequence 120, App
807	5	1.2	177	1	US-08-850-118-4	Sequence 4, Appli
808	5	1.2	177	2	US-09-008-253-4	Sequence 4, Appli
809	5	1.2	177	3	US-09-093-335-4	Sequence 4, Appli
810	5	1.2	180	1	US-08-844-064-2	Sequence 2, Appli
811	5	1.2	180	1	US-08-447-010-6	Sequence 6, Appli
812	5	1.2	180	1	US-08-447-010-7	Sequence 7, Appli
813	5	1.2	180	1	US-08-447-010-8	Sequence 8, Appli
814	5	1.2	180	3	US-09-009-433-2	Sequence 2, Appli
815	5	1.2	182	1	US-07-991-867B-10	Sequence 10, Appl
816	5	1.2	182	1	US-08-107-755A-10	Sequence 10, Appl
817	5	1.2	182	1	US-08-226-264-28	Sequence 28, Appl
818	5	1.2	182	2	US-08-544-332-10	Sequence 10, Appl
819	5	1.2	183	2	US-08-980-329C-2	Sequence 2, Appli
820	5	1.2	184	1	US-08-044-621D-30	Sequence 30, Appl
821	5	1.2	184	1	US-08-044-621D-31	Sequence 31, Appl
822	5	1.2	184	1	US-08-709-912-1	Sequence 1, Appli
823	5	1.2	184	2	US-08-531-525-37	Sequence 37, Appl
824	5	1.2	184	2	US-09-047-370-1	Sequence 1, Appli
825	5	1.2	184	2	US-08-718-270A-37	Sequence 37, Appl
826	5	1.2	184	4	US-09-269-258A-2	Sequence 2, Appli
827	5	1.2	185	1	US-08-709-912-2	Sequence 2, Appli
828	5	1.2	185	2	US-09-047-370-2	Sequence 2, Appli
829	5	1.2	185	2	US-08-933-750C-2	Sequence 2, Appli
830	5	1.2	185	3	US-09-045-632-26	Sequence 26, Appl

831	5	1.2	185	3	US-08-965-689A-1	Sequence 1, Appli
832	5	1.2	185	4	US-09-234-613-2	Sequence 2, Appli
833	5	1.2	185	4	US-09-359-967-1	Sequence 1, Appli
834	5	1.2	186	1	US-08-117-083-22	Sequence 22, Appl
835	5	1.2	186	2	US-08-750-194-2	Sequence 2, Appli
836	5	1.2	186	2	US-08-557-309B-43	Sequence 43, Appl
837	5	1.2	186	3	US-08-834-306-43	Sequence 43, Appl
838	5	1.2	186	4	US-08-993-674A-43	Sequence 43, Appl
839	5	1.2	187	1	US-08-049-473-33	Sequence 33, Appl
840	5	1.2	187	1	US-08-312-648-33	Sequence 33, Appl
841	5	1.2	187	5	PCT-US94-04190-33	Sequence 33, Appl
842	5	1.2	188	2	US-08-905-267-2	Sequence 2, Appli
843	5	1.2	188	3	US-09-332-934-2	Sequence 2, Appli
844	5	1.2	188	4	US-09-314-051-2	Sequence 2, Appli
845	5	1.2	190	1	US-07-956-700B-15	Sequence 15, Appl
846	5	1.2	190	1	US-07-956-700B-30	Sequence 30, Appl
847	5	1.2	190	1	US-08-476-537-15	Sequence 15, Appl
848	5	1.2	190	1	US-08-476-537-30	Sequence 30, Appl
849	5	1.2	190	1	US-08-485-607-15	Sequence 15, Appl
850	5	1.2	190	1	US-08-485-607-30	Sequence 30, Appl
851	5	1.2	190	2	US-08-475-879-15	Sequence 15, Appl
852	5	1.2	190	2	US-08-475-879-30	Sequence 30, Appl
853	5	1.2	191	5	PCT-US95-07171-3	Sequence 3, Appli
854	5	1.2	192	4	US-09-475-316A-23	Sequence 23, Appl
855	5	1.2	192	4	US-09-475-316A-33	Sequence 33, Appl
856	5	1.2	192	4	US-09-475-316A-35	Sequence 35, Appl
857	5	1.2	193	2	US-08-896-605A-2	Sequence 2, Appli
858	5	1.2	193	2	US-08-896-501A-2	Sequence 2, Appli
859	5	1.2	193	4	US-08-832-180-9	Sequence 9, Appli
860	5	1.2	194	1	US-08-118-469A-7	Sequence 7, Appli
861	5	1.2	194	1	US-08-909-119-7	Sequence 7, Appli
862	5	1.2	194	3	US-08-480-640A-196	Sequence 196, App
863	5	1.2	194	4	US-08-686-968C-196	Sequence 196, App
864	5	1.2	194	4	US-08-488-237A-196	Sequence 196, App
865	5	1.2	195	2	US-08-822-260-1	Sequence 1, Appli
866	5	1.2	196	1	US-08-188-582-22	Sequence 22, Appl
867	5	1.2	196	1	US-08-646-715-22	Sequence 22, Appl
868	5	1.2	196	2	US-08-829-110-5	Sequence 5, Appli
869	5	1.2	196	2	US-08-078-311-13	Sequence 13, Appl
870	5	1.2	196	2	US-08-748-483-3	Sequence 3, Appli
871	5	1.2	196	2	US-08-460-402-13	Sequence 13, Appl
872	5	1.2	196	5	PCT-US91-09055-5	Sequence 5, Appli
873	5	1.2	197	2	US-08-377-309-3	Sequence 3, Appli
874	5	1.2	197	5	PCT-US96-00996-6	Sequence 6, Appli
875	5	1.2	200	1	US-08-652-859-3	Sequence 3, Appli
876	5	1.2	200	2	US-08-919-706-3	Sequence 3, Appli
877	5	1.2	200	2	US-09-153-751-3	Sequence 3, Appli
878	5	1.2	200	4	US-09-181-183-38	Sequence 38, Appl
879	5	1.2	201	1	US-08-155-171B-19	Sequence 19, Appl
880	5	1.2	201	1	US-08-415-751-2	Sequence 2, Appli
881	5	1.2	201	2	US-08-435-998-19	Sequence 19, Appl
882	5	1.2	203	2	US-08-598-873-4	Sequence 4, Appli
883	5	1.2	203	4	US-08-605-430-4	Sequence 4, Appli
884	5	1.2	204	1	US-08-155-171B-20	Sequence 20, Appl
885	5	1.2	204	2	US-08-531-525-48	Sequence 48, Appl
886	5	1.2	204	2	US-08-718-270A-48	Sequence 48, Appl
887	5	1.2	204	2	US-08-435-998-20	Sequence 20, Appl
888	5	1.2	205	1	US-08-321-162-2	Sequence 2, Appli
889	5	1.2	205	1	US-08-448-736-1	Sequence 1, Appli
890	5	1.2	205	1	US-08-441-216-2	Sequence 2, Appli
891	5	1.2	205	1	US-08-452-779-1	Sequence 1, Appli
892	5	1.2	205	1	US-08-299-567-4	Sequence 4, Appli
893	5	1.2	205	2	US-08-445-065-1	Sequence 1, Appli
894	5	1.2	205	2	US-08-445-065-12	Sequence 12, Appl
895	5	1.2	205	3	US-08-959-524-1	Sequence 1, Appli
896	5	1.2	205	3	US-08-959-524-12	Sequence 12, Appl
897	5	1.2	206	6	5189147-9	Patent No. 5189147
898	5	1.2	208	2	US-08-766-551-9	Sequence 9, Appli
899	5	1.2	208	4	US-08-961-083-214	Sequence 214, App
900	5	1.2	209	4	US-09-071-709-11	Sequence 11, Appl
901	5	1.2	210	2	US-08-969-106-13	Sequence 13, Appl
902	5	1.2	211	1	US-08-244-686-2	Sequence 2, Appli
903	5	1.2	212	1	US-08-915-003-9	Sequence 9, Appli
904	5	1.2	212	2	US-08-642-247-9	Sequence 9, Appli
905	5	1.2	213	2	US-08-941-263-3	Sequence 3, Appli
906	5	1.2	213	3	US-09-186-250-2	Sequence 2, Appli
907	5	1.2	213	3	US-09-186-250-6	Sequence 6, Appli
908	5	1.2	213	3	US-09-227-178-3	Sequence 3, Appli
909	5	1.2	213	4	US-09-470-449-3	Sequence 3, Appli
910	5	1.2	213	4	US-09-517-347-2	Sequence 2, Appli
911	5	1.2	213	4	US-09-517-347-6	Sequence 6, Appli
912	5	1.2	215	2	US-08-796-676-3	Sequence 3, Appli
913	5	1.2	215	2	US-08-967-364-8	Sequence 8, Appli
914	5	1.2	215	3	US-09-368-408-8	Sequence 8, Appli
915	5	1.2	216	2	US-08-840-683-9	Sequence 9, Appli
916	5	1.2	216	2	US-08-737-716-4	Sequence 4, Appli
917	5	1.2	216	2	US-08-555-722-9	Sequence 9, Appli
918	5	1.2	217	2	US-08-176-414B-3	Sequence 3, Appli
919	5	1.2	217	6	5189147-7	Patent No. 5189147
920	5	1.2	218	1	US-08-126-593A-4	Sequence 4, Appli
921	5	1.2	218	1	US-08-454-039A-4	Sequence 4, Appli
922	5	1.2	218	1	US-08-470-837-26	Sequence 26, Appl
923	5	1.2	219	1	US-08-461-005-13	Sequence 13, Appl
924	5	1.2	219	2	US-08-855-140-4	Sequence 4, Appli
925	5	1.2	219	2	US-08-796-676-1	Sequence 1, Appli
926	5	1.2	219	2	US-07-930-685-13	Sequence 13, Appl
927	5	1.2	219	2	US-08-850-880-5	Sequence 5, Appli
928	5	1.2	219	2	US-08-944-916-5	Sequence 5, Appli
929	5	1.2	219	2	US-08-814-877-5	Sequence 5, Appli
930	5	1.2	219	4	US-08-928-213B-60	Sequence 60, Appl
931	5	1.2	221	2	US-08-916-901-5	Sequence 5, Appli
932	5	1.2	221	3	US-09-081-320-5	Sequence 5, Appli
933	5	1.2	221	3	US-09-081-320-27	Sequence 27, Appl
934	5	1.2	222	2	US-08-773-423-1	Sequence 1, Appli
935	5	1.2	222	3	US-08-942-001-8	Sequence 8, Appli
936	5	1.2	222	4	US-09-181-183-28	Sequence 28, Appl
937	5	1.2	222	4	US-09-337-386-8	Sequence 8, Appli
938	5	1.2	223	1	US-07-708-885B-3	Sequence 3, Appli
939	5	1.2	223	1	US-07-714-386-3	Sequence 3, Appli
940	5	1.2	223	1	US-07-708-888A-3	Sequence 3, Appli
941	5	1.2	223	2	US-08-820-170A-16	Sequence 16, Appl
942	5	1.2	223	3	US-09-055-699-16	Sequence 16, Appl
943	5	1.2	223	4	US-09-273-565-16	Sequence 16, Appl
944	5	1.2	224	4	US-09-091-899-10	Sequence 10, Appl
945	5	1.2	224	4	US-09-045-193-2	Sequence 2, Appli
946	5	1.2	225	1	US-07-708-885B-1	Sequence 1, Appli
947	5	1.2	225	1	US-07-714-386-1	Sequence 1, Appli
948	5	1.2	225	1	US-07-708-888A-1	Sequence 1, Appli
949	5	1.2	225	4	US-09-181-183-30	Sequence 30, Appl
950	5	1.2	226	1	US-08-378-011A-3	Sequence 3, Appli
951	5	1.2	226	4	US-08-896-933-24	Sequence 24, Appl
952	5	1.2	226	5	PCT-US96-10602-14	Sequence 14, Appl
953	5	1.2	227	1	US-08-254-493-1	Sequence 1, Appli
954	5	1.2	227	2	US-08-253-751-6	Sequence 6, Appli
955	5	1.2	227	2	US-08-453-925-6	Sequence 6, Appli
956	5	1.2	228	1	US-08-447-591-2	Sequence 2, Appli
957	5	1.2	228	1	US-08-447-591-3	Sequence 3, Appli
958	5	1.2	228	1	US-08-447-591-4	Sequence 4, Appli
959	5	1.2	228	1	US-08-450-943-2	Sequence 2, Appli
960	5	1.2	228	1	US-08-450-943-3	Sequence 3, Appli
961	5	1.2	228	1	US-08-450-943-4	Sequence 4, Appli
962	5	1.2	228	1	US-08-059-031-2	Sequence 2, Appli
963	5	1.2	228	1	US-08-059-031-3	Sequence 3, Appli
964	5	1.2	228	1	US-08-059-031-4	Sequence 4, Appli
965	5	1.2	228	1	US-08-408-222B-1	Sequence 1, Appli
966	5	1.2	228	2	US-08-484-101B-40	Sequence 40, Appl
967	5	1.2	228	2	US-08-417-495-27	Sequence 27, Appl
968	5	1.2	228	2	US-08-284-391B-27	Sequence 27, Appl
969	5	1.2	228	2	US-08-450-942-2	Sequence 2, Appli
970	5	1.2	228	2	US-08-450-942-3	Sequence 3, Appli
971	5	1.2	228	2	US-08-450-942-4	Sequence 4, Appli
972	5	1.2	228	5	PCT-US92-01785-27	Sequence 27, Appl
973	5	1.2	228	5	PCT-US94-05090-2	Sequence 2, Appli
974	5	1.2	228	5	PCT-US94-05090-3	Sequence 3, Appli
975	5	1.2	228	5	PCT-US94-05090-4	Sequence 4, Appli
976	5	1.2	228	5	PCT-US95-00454-27	Sequence 27, Appl

977	5	1.2	229	3	US-09-024-023-2	Sequence 2, Appli
978	5	1.2	229	5	PCT-US96-03916-13	Sequence 13, Appl
979	5	1.2	229	5	PCT-US96-03916-62	Sequence 62, Appl
980	5	1.2	231	2	US-08-808-550-33	Sequence 33, Appl
981	5	1.2	231	4	US-08-448-489-19	Sequence 19, Appl
982	5	1.2	232	5	PCT-US94-09700-6	Sequence 6, Appli
983	5	1.2	233	1	US-08-446-918A-4	Sequence 4, Appli
984	5	1.2	233	2	US-08-557-309B-40	Sequence 40, Appl
985	5	1.2	233	2	US-08-580-806-4	Sequence 4, Appli
986	5	1.2	233	3	US-08-834-306-40	Sequence 40, Appl
987	5	1.2	233	4	US-08-993-674A-40	Sequence 40, Appl
988	5	1.2	234	1	US-08-684-862-6	Sequence 6, Appli
989	5	1.2	234	1	US-08-850-118-2	Sequence 2, Appli
990	5	1.2	234	2	US-09-008-253-2	Sequence 2, Appli
991	5	1.2	234	3	US-09-093-335-2	Sequence 2, Appli
992	5	1.2	235	3	US-08-807-151-3	Sequence 3, Appli
993	5	1.2	236	1	US-08-378-011A-1	Sequence 1, Appli
994	5	1.2	236	2	US-08-738-413B-10	Sequence 10, Appl
995	5	1.2	238	4	US-09-071-709-3	Sequence 3, Appli
996	5	1.2	238	4	US-08-896-933-28	Sequence 28, Appl
997	5	1.2	239	2	US-08-860-174A-2	Sequence 2, Appli
998	5	1.2	239	3	US-08-462-436-2	Sequence 2, Appli
999	5	1.2	239	4	US-08-465-275-2	Sequence 2, Appli
1000	5	1.2	240	1	US-08-261-822A-79	Sequence 79, Appl

ALIGNMENTS

```
RESULT   1
US-09-335-409-15
; Sequence 15, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
;  APPLICANT: Schupp, Thomas
;  APPLICANT: Ligon, James
;  APPLICANT: Molnar, Istvan
;  APPLICANT: Zirkle, Ross
;  APPLICANT: Cyr, Devon
;  APPLICANT: Goerlach, Joern
;  TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
;  FILE REFERENCE: 4-30582A
;  CURRENT APPLICATION NUMBER: US/09/335,409
;  CURRENT FILING DATE: 1999-06-17
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;    LENGTH: 145
;    TYPE: PRT
;    ORGANISM: Sorangium cellulosum
US-09-335-409-15


  Query Match             1.6%;  Score 7;  DB 3;  Length 145;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      83 SVRPCIR 89
           |||||||
Db     127 SVRPCIR 133


RESULT   2
US-09-053-866-2
; Sequence 2, Application US/09053866
; Patent No. 6111075
;  GENERAL INFORMATION:
;    APPLICANT:  Xu, Wenfeng
;    APPLICANT:  Presnell, Scott R.
;    APPLICANT:  Yee, David P.
;    APPLICANT:  Foster, Donald C.
;    TITLE OF INVENTION:  PROTEASE-ACTIVATED RECEPTOR
;    TITLE OF INVENTION:  PAR4 (ZCHEMR2)
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  1201 Eastlake Avenue East
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/053,866
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leith, Debra K
;      REGISTRATION NUMBER:  32,619
;      REFERENCE/DOCKET NUMBER:  98-10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-442-6674
;      TELEFAX:  206-442-6678
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  385 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-09-053-866-2


  Query Match             1.6%;  Score 7;  DB 3;  Length 385;
  Best Local Similarity   100.0%;  Pred. No. 47;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     380 PSLALST 386
           |||||||
Db     324 PSLALST 330


RESULT   3
US-07-730-953-2
; Sequence 2, Application US/07730953
; Patent No. 5288614
;  GENERAL INFORMATION:
;    APPLICANT:  BODENMULLER, Heinz
;    APPLICANT:  DESSAUER, Andreas
;    TITLE OF INVENTION:  METHOD FOR THE DETECTION OF MALIGNANT
;    TITLE OF INVENTION:  DISEASES
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/730,953
;      FILING DATE:  19910723
```

```
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 39 42 999.7
;      FILING DATE:  21-DEC-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kitts, Monica C.
;      REGISTRATION NUMBER:  36,105
;      REFERENCE/DOCKET NUMBER:  P564-1119
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  400 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-730-953-2


  Query Match             1.6%;  Score 7;  DB 1;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     106 GLKEELA 112
           |||||||
Db     206 GLKEELA 212


RESULT   4
US-09-001-984C-106
; Sequence 106, Application US/09001984C
; Patent No. 6245331
; GENERAL INFORMATION:
;  APPLICANT:  Laal, Suman
;  APPLICANT:  Zolla-Pazner, Susan
;  APPLICANT:  Belisle, John T
;  TITLE OF INVENTION:  EARLY DETECTION OF MYCOBACTERIAL DISEASE
;  FILE REFERENCE:  NYU-011
;  CURRENT APPLICATION NUMBER: US/09/001,984C
;  CURRENT FILING DATE:  1997-12-31
;  PRIOR APPLICATION NUMBER:  60/034,003
;  PRIOR FILING DATE:  1996-12-31
;  NUMBER OF SEQ ID NOS:  106
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 106
;   LENGTH: 741
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis
US-09-001-984C-106


  Query Match             1.6%;  Score 7;  DB 4;  Length 741;
  Best Local Similarity   100.0%;  Pred. No. 83;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     336 TGSIYIV 342
           |||||||
Db     390 TGSIYIV 396


RESULT   5
US-09-208-742-4
; Sequence 4, Application US/09208742
; Patent No. 6174679
; GENERAL INFORMATION:
;  APPLICANT: Kaufmann, Joerg
;  TITLE OF INVENTION: CIF150/hTAFII150 is Necessary for Cell
;  TITLE OF INVENTION: Cycle Progression
;  FILE REFERENCE: 1453.002
;  CURRENT APPLICATION NUMBER: US/09/208,742
;  CURRENT FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 1261
;   TYPE: PRT
;   ORGANISM: human
US-09-208-742-4


  Query Match             1.6%;  Score 7;  DB 4;  Length 1261;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     350 PSKWNWL 356
           |||||||
Db     527 PSKWNWL 533


RESULT   6
US-08-447-411-71
; Sequence 71, Application US/08447411
; Patent No. 5773243
;  GENERAL INFORMATION:
;    APPLICANT:  FRITZINGER, DAVID C.
;    APPLICANT:  BREDEHORST, REINHARD
;    APPLICANT:  VOGEL, CARL-WILHELM
;    TITLE OF INVENTION:  DNA ENCODING COBRA C3, CVF1, AND CVF2
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,411
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/043,747
;      FILING DATE:  07-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5773243man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-101-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  71:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
\    US-08-447-411-71


  Query Match             1.4%;  Score 6;  DB 1;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 39;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     201 SDFPKS 206
           ||||||
Db      17 SDFPKS 22
```

```
RESULT   7
US-08-662-227-29
; Sequence 29, Application US/08662227
; Patent No. 5922320
;  GENERAL INFORMATION:
;    APPLICANT:  VOGEL, CARL-WILHELM
;    APPLICANT:  BREDEHORST, REINHORST
;    APPLICANT:  KOCK, MICHAEL
;    APPLICANT:  FRITZINGER, DAVID
;    TITLE OF INVENTION:  RECOMBINANT PROCVF
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/662,227
;      FILING DATE:  14-JUN-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-0107-0X
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-662-227-29


  Query Match             1.4%;  Score 6;  DB 2;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 39;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     201 SDFPKS 206
           ||||||
Db      17 SDFPKS 22


RESULT   8
US-08-480-473B-47
; Sequence 47, Application US/08480473B
; Patent No. 5882914
;  GENERAL INFORMATION:
;    APPLICANT:  Semenza, Gregg L.
;    TITLE OF INVENTION:  HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
;    NUMBER OF SEQUENCES:  64
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,473B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07265/053001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619/678-5070
;      TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  102 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-480-473B-47


  Query Match             1.4%;  Score 6;  DB 2;  Length 102;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     147 MIYISD 152
           ||||||
Db       4 MIYISD 9


RESULT   9
US-08-915-213-47
; Sequence 47, Application US/08915213
; Patent No. 6020462
;  GENERAL INFORMATION:
;    APPLICANT:  Semenza, Gregg L.
;    TITLE OF INVENTION:  HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
;    NUMBER OF SEQUENCES:  64
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/915,213
;      FILING DATE:  20-AUG-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/480,473
;      FILING DATE:  06-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07265/053001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619/678-5070
;      TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  102 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-915-213-47
```

```
  Query Match             1.4%;  Score 6;  DB 3;  Length 102;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    147 MIYISD 152
          ||||||
Db      4 MIYISD 9


RESULT  10
US-09-235-217-47
; Sequence 47, Application US/09235217
; Patent No. 6222018
;  GENERAL INFORMATION:
;    APPLICANT:  Semenza, Gregg L.
;    TITLE OF INVENTION:  HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
;    NUMBER OF SEQUENCES:  64
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/235,217
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/480,473
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07265/053001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619/678-5070
;      TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  102 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-235-217-47


  Query Match             1.4%;  Score 6;  DB 4;  Length 102;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    147 MIYISD 152
          ||||||
Db      4 MIYISD 9


RESULT  11
US-08-466-583-5
; Sequence 5, Application US/08466583
; Patent No. 5919998
;  GENERAL INFORMATION:
;    APPLICANT:  Bandurski, Robert S.
;    APPLICANT:  Szerszen, Jedrzej B.
;    APPLICANT:  Szczyglowski, Krzysztof
;    TITLE OF INVENTION:  Genetic Control of Plant Hormone Levels
;    TITLE OF INVENTION:  and Plant Growth.
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee & Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  United States of America
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/466,583
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/265,427
;      FILING DATE:  24-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  11-94A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  303/499-8080
;      TELEFAX:  303/499-8089
;      TELEX:  49617824
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  131 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
US-08-466-583-5


  Query Match             1.4%;  Score 6;  DB 2;  Length 131;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    108 KEELAR 113
          ||||||
Db     29 KEELAR 34


RESULT  12
PCT-US95-07820-5
; Sequence 5, Application PC/TUS9507820
;  GENERAL INFORMATION:
;    APPLICANT:  Bandurski, Robert S.
;    APPLICANT:  Szerszen, Jedrzej B.
;    APPLICANT:  Szczyglowski, Krzysztof
;    TITLE OF INVENTION:  Genetic Control of Plant Hormone Levels
;    TITLE OF INVENTION:  and Plant Growth.
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee & Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  United States of America
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  PCT/US95/07820
;      FILING DATE:  19-JUN-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/265,427
;      FILING DATE:  24-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  11-94B PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  303/499-8080
;      TELEFAX:  303/499-8089
;      TELEX:  49617824
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  131 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
PCT-US95-07820-5


  Query Match             1.4%;  Score 6;  DB 5;  Length 131;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     108 KEELAR 113
           ||||||
Db      29 KEELAR 34


RESULT  13
US-08-240-712-21
; Sequence 21, Application US/08240712
; Patent No. 5599907
;  GENERAL INFORMATION:
;    APPLICANT:  ANDERSON, DAVID C.
;    APPLICANT:  MATHEWS, ANTONY JAMES
;    APPLICANT:  STETLER, GARY L.
;    TITLE OF INVENTION:  PRODUCTION AND USE OF MULTIMERIC
;    TITLE OF INVENTION:  HEMOGLOBINS
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/240,712
;      FILING DATE:  09-MAY-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/09752
;      FILING DATE:  13-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOPER, IVER P
;      REGISTRATION NUMBER:  28,005
;      REFERENCE/DOCKET NUMBER:  ANDERSON=6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  146 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-240-712-21


  Query Match             1.4%;  Score 6;  DB 1;  Length 146;
  Best Local Similarity   100.0%;  Pred. No. 2e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     307 KVLTSL 312
           ||||||
Db      66 KVLTSL 71


RESULT  14
US-08-240-712-22
; Sequence 22, Application US/08240712
; Patent No. 5599907
;  GENERAL INFORMATION:
;    APPLICANT:  ANDERSON, DAVID C.
;    APPLICANT:  MATHEWS, ANTONY JAMES
;    APPLICANT:  STETLER, GARY L.
;    TITLE OF INVENTION:  PRODUCTION AND USE OF MULTIMERIC
;    TITLE OF INVENTION:  HEMOGLOBINS
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/240,712
;      FILING DATE:  09-MAY-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/09752
;      FILING DATE:  13-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOPER, IVER P
;      REGISTRATION NUMBER:  28,005
;      REFERENCE/DOCKET NUMBER:  ANDERSON=6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  146 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-240-712-22


  Query Match             1.4%;  Score 6;  DB 1;  Length 146;
  Best Local Similarity   100.0%;  Pred. No. 2e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     307 KVLTSL 312
           ||||||
```

```
Db      66 KVLTSL 71


RESULT  15
US-08-240-712-23
; Sequence 23, Application US/08240712
; Patent No. 5599907
;  GENERAL INFORMATION:
;    APPLICANT:  ANDERSON, DAVID C.
;    APPLICANT:  MATHEWS, ANTONY JAMES
;    APPLICANT:  STETLER, GARY L.
;    TITLE OF INVENTION:  PRODUCTION AND USE OF MULTIMERIC
;    TITLE OF INVENTION:  HEMOGLOBINS
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/240,712
;      FILING DATE:  09-MAY-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/09752
;      FILING DATE:  13-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOPER, IVER P
;      REGISTRATION NUMBER:  28,005
;      REFERENCE/DOCKET NUMBER:  ANDERSON=6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  146 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-240-712-23


  Query Match             1.4%;  Score 6;  DB 1;  Length 146;
  Best Local Similarity   100.0%;  Pred. No. 2e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     307 KVLTSL 312
           ||||||
Db      66 KVLTSL 71


Search completed: October 30, 2001, 15:30:12
Job time: 303 sec
```

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 15:26:24 ; Search time 17.77 Seconds
(without alignments)
1826.132 Million cell updates/sec

Title: US-09-619-049-855
Perfect score: 426
Sequence: 1 MRLPWVIFCTVLLLIFTNNS.....NKFTNEEEGRINCINFDSQT 426

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	8	1.9	266	2	A48349	UL28 protein - ate
2	8	1.9	664	1	WMBEBH	infected cell prot
3	8	1.9	724	1	WMBEPR	infected cell prot
4	8	1.9	766	1	WZBEC5	infected cell prot
5	8	1.9	773	2	T42575	infected cell prot
6	8	1.9	780	1	WMBEH8	infected cell prot
7	8	1.9	785	1	WMBEK8	infected cell prot
8	8	1.9	785	1	WMBEW8	infected cell prot
9	8	1.9	3738	2	T05501	hypothetical prote
10	8	1.9	3795	2	T00831	hypothetical prote
11	7	1.6	74	2	T14687	hypothetical prote
12	7	1.6	76	2	C84094	protein translocat
13	7	1.6	122	1	M3HUGA	Ig heavy chain V-I
14	7	1.6	137	2	E82653	hypothetical prote
15	7	1.6	202	2	S25660	4-hydroxybenzoate
16	7	1.6	202	2	B86097	chorismate lyase [
17	7	1.6	214	2	T14920	hypothetical prote
18	7	1.6	262	1	A26324	deoxyribonuclease
19	7	1.6	307	2	H64541	fructose-bisphosph
20	7	1.6	311	2	T01154	hypothetical prote
21	7	1.6	328	2	D70574	hypothetical prote
22	7	1.6	344	2	S57901	estradiol 17beta-d
23	7	1.6	344	2	S62652	estradiol 17beta-d
24	7	1.6	353	1	D65025	hypothetical prote
25	7	1.6	353	2	B72092	probable endonucle
26	7	1.6	353	2	A86532	phopholipase D typ
27	7	1.6	353	2	G85892	probable permease
28	7	1.6	367	2	T23825	hypothetical prote
29	7	1.6	372	2	C70590	probable nicT prot
30	7	1.6	380	2	B69523	succinyl-CoA synth
31	7	1.6	380	2	E69319	flavoprotein reduc
32	7	1.6	399	2	A25470	cytokeratin 19 - b
33	7	1.6	400	1	KRHU9	keratin 19, type I
34	7	1.6	401	2	S57657	keratin 19 - potor
35	7	1.6	401	2	D71003	probable transamin
36	7	1.6	430	2	A43285	cyclin B1 - mouse
37	7	1.6	430	2	I48316	cyclin B1 - mouse
38	7	1.6	473	2	JC4313	keratin 16, type I
39	7	1.6	473	4	A33652	probable keratin 1
40	7	1.6	480	2	H71430	hypothetical prote
41	7	1.6	521	2	D75581	malate synthase -
42	7	1.6	533	2	S42404	thromboxane-A synt
43	7	1.6	554	2	T50118	kinesin-related pr
44	7	1.6	651	2	B86026	hypothetical prote
45	7	1.6	651	2	S47750	hypothetical prote
46	7	1.6	731	2	T44752	probable malate sy
47	7	1.6	741	2	F70722	probable glcB prot
48	7	1.6	743	2	T32421	hypothetical prote
49	7	1.6	791	2	E71333	hypothetical prote
50	7	1.6	792	2	T21276	hypothetical prote
51	7	1.6	819	2	T08745	probable RNA helic
52	7	1.6	872	2	S51620	Ost oncogene - rat
53	7	1.6	1206	2	T44376	DNA-directed RNA p
54	7	1.6	1273	2	S58782	SEC31 protein - ye
55	7	1.6	1336	2	T17479	hypothetical prote
56	7	1.6	1467	2	T48162	hypothetical prote
57	7	1.6	2010	2	B71616	phosphatase (acid
58	7	1.6	2241	2	T16064	hypothetical prote
59	7	1.6	3085	2	T00327	polyprotein - infe
60	7	1.6	3225	2	D81702	adherence factor T
61	6	1.4	27	2	B37043	Ig heavy chain, po
62	6	1.4	30	2	E84786	hypothetical prote
63	6	1.4	38	2	S78357	photosystem II pro
64	6	1.4	42	2	G56271	lpfA 5'-region hyp
65	6	1.4	47	2	E70227	hypothetical prote
66	6	1.4	56	2	I40011	sspF protein - Bac
67	6	1.4	67	2	E81925	probable sec-indep
68	6	1.4	69	2	S40540	gef protein - Esch
69	6	1.4	69	2	H85481	hypothetical prote
70	6	1.4	69	2	T33062	hypothetical prote
71	6	1.4	72	2	B34669	calmodulin-related
72	6	1.4	76	1	A70028	hypothetical prote
73	6	1.4	76	2	T24688	hypothetical prote
74	6	1.4	80	2	C69204	hypothetical prote
75	6	1.4	84	2	B84820	hypothetical prote
76	6	1.4	84	2	B48652	hypothetical prote
77	6	1.4	87	2	F49754	holin component xh
78	6	1.4	87	2	T09905	hypothetical prote
79	6	1.4	92	2	JC1208	DNA-binding protei
80	6	1.4	92	2	S00015	DNA-binding protei
81	6	1.4	92	2	G86886	30S ribosomal prot
82	6	1.4	100	2	JH0269	cystatin - avocado
83	6	1.4	102	2	S22356	hemoglobin beta-z
84	6	1.4	105	2	B54864	mannose/glucose-sp
85	6	1.4	106	2	T18157	hypothetical prote
86	6	1.4	106	2	T18186	hypothetical prote
87	6	1.4	108	2	A59096	hypothetical prote
88	6	1.4	108	2	G64386	hypothetical prote
89	6	1.4	109	2	T00234	hypothetical prote
90	6	1.4	113	2	A72387	hypothetical prote
91	6	1.4	115	2	A72501	hypothetical prote
92	6	1.4	115	2	D71541	probable inclusion
93	6	1.4	116	2	G72686	hypothetical prote
94	6	1.4	117	2	H64094	ribosomal protein
95	6	1.4	117	2	A56611	HTF4a 5'-region hy
96	6	1.4	118	2	C84988	30S ribosomal prot
97	6	1.4	118	2	D82061	ribosomal protein
98	6	1.4	122	2	T17991	hypothetical prote
99	6	1.4	123	2	G82357	conserved hypothet
100	6	1.4	123	2	H81694	inclusion membrane
101	6	1.4	124	2	T37750	very hypothetical
102	6	1.4	129	2	T04230	hypothetical prote

103	6	1.4	133	2	T05440	hypothetical prote
104	6	1.4	134	2	A54002	pollen allergen Lo
105	6	1.4	136	2	B85685	hypothetical prote
106	6	1.4	140	2	S36804	profilin II - huma
107	6	1.4	142	1	BVECTR	tolR protein - Esc
108	6	1.4	142	2	F85576	hypothetical prote
109	6	1.4	145	2	PC1247	Sin a I allergen 2
110	6	1.4	145	2	D86904	conserved hypothet
111	6	1.4	146	1	HGMQJ	hemoglobin gamma c
112	6	1.4	146	1	HGBAY	hemoglobin gamma c
113	6	1.4	146	1	HGMQP	hemoglobin gamma c
114	6	1.4	147	1	HGHUA	hemoglobin gamma-A
115	6	1.4	147	1	HGCZA	hemoglobin gamma-A
116	6	1.4	147	1	HGHUG	hemoglobin gamma-G
117	6	1.4	147	1	HGCZG	hemoglobin gamma-G
118	6	1.4	147	1	I37025	hemoglobin gamma-G
119	6	1.4	147	1	I37022	hemoglobin gamma-A
120	6	1.4	147	1	HGMQR	hemoglobin gamma-1
121	6	1.4	147	1	HGMQR2	hemoglobin gamma-2
122	6	1.4	147	1	HGMKS	hemoglobin gamma c
123	6	1.4	147	1	HBMSH0	hemoglobin beta-h0
124	6	1.4	147	1	HBMSH1	hemoglobin beta-h1
125	6	1.4	147	2	B27800	hemoglobin gamma-2
126	6	1.4	147	2	I37002	hemoglobin gamma c
127	6	1.4	147	2	I77303	hemoglobin gamma c
128	6	1.4	147	2	I77304	hemoglobin gamma c
129	6	1.4	147	2	I37035	hemoglobin gamma-1
130	6	1.4	147	2	I37036	hemoglobin gamma-2
131	6	1.4	147	2	G00014	hemoglobin gamma-1
132	6	1.4	147	2	S22357	hemoglobin beta-y
133	6	1.4	147	2	G00015	hemoglobin gamma-2
134	6	1.4	147	2	A27800	hemoglobin gamma-1
135	6	1.4	149	2	S58314	calmodulin - moss
136	6	1.4	149	2	JC5097	ferric uptake regu
137	6	1.4	151	1	MCDO	calmodulin - slime
138	6	1.4	151	2	JN0722	calmodulin - Pneum
139	6	1.4	151	2	T27111	hypothetical prote
140	6	1.4	155	4	S21403	hemoglobin gamma-G
141	6	1.4	156	2	T49529	hypothetical prote
142	6	1.4	159	1	S73354	conserved hypothet
143	6	1.4	163	2	F72728	probable ribosomal
144	6	1.4	167	2	T39264	hypothetical prote
145	6	1.4	168	2	S18374	flavodoxin - Anaba
146	6	1.4	168	2	T29679	hypothetical prote
147	6	1.4	168	2	G83674	phenylacetic acid
148	6	1.4	169	1	F70157	nucleoside-diphosp
149	6	1.4	169	2	D72347	hypothetical prote
150	6	1.4	170	1	S04600	flavodoxin - Anaba
151	6	1.4	170	2	S64488	regulatory protein
152	6	1.4	170	2	F72115	hypothetical prote
153	6	1.4	170	2	F86506	hypothetical prote
154	6	1.4	171	2	T36980	hypothetical prote
155	6	1.4	171	2	D70832	hypothetical prote
156	6	1.4	172	2	A71455	probable acetyltra
157	6	1.4	175	2	S69636	hypothetical prote
158	6	1.4	176	2	T48292	hypothetical prote
159	6	1.4	178	2	A64115	parB protein homol
160	6	1.4	179	2	A48474	ICP18.5 homolog -
161	6	1.4	180	2	A32246	70K antigen - Chla
162	6	1.4	181	2	S69721	probable adenine p
163	6	1.4	186	2	S75396	hypothetical prote
164	6	1.4	186	2	T21187	hypothetical prote
165	6	1.4	187	2	T02887	probable calmoduli
166	6	1.4	188	2	S32862	outH protein - Erw
167	6	1.4	188	2	A84383	probable acetyltra
168	6	1.4	193	2	T44148	hypothetical prote
169	6	1.4	193	2	A40738	surface antigen CD
170	6	1.4	194	2	T26391	hypothetical prote
171	6	1.4	195	2	T39667	probable antioxida
172	6	1.4	197	2	C83003	imidazoleglycerol-
173	6	1.4	198	2	S65530	sodium-translocati
174	6	1.4	200	2	T44143	DR3 protein [impor
175	6	1.4	200	2	T43953	hypothetical prote
176	6	1.4	201	2	C55331	guanylate cyclase
177	6	1.4	202	2	B55331	guanylate cyclase-
178	6	1.4	202	2	T24134	hypothetical prote
179	6	1.4	203	2	D82993	conserved hypothet
180	6	1.4	206	2	H84165	hypothetical prote
181	6	1.4	207	2	D64348	hypothetical prote
182	6	1.4	213	2	T18684	hypothetical prote
183	6	1.4	213	2	A86654	hypothetical prote
184	6	1.4	213	2	S27736	hypothetical prote
185	6	1.4	215	2	H64849	yceH protein - Esc
186	6	1.4	215	2	A85669	hypothetical prote
187	6	1.4	216	2	B69253	hypothetical prote
188	6	1.4	216	2	T24080	hypothetical prote
189	6	1.4	218	2	S75100	ABC transport prot
190	6	1.4	222	2	T03613	GTP-binding protei
191	6	1.4	222	2	T03622	GTP-binding protei
192	6	1.4	222	2	T31541	hypothetical prote
193	6	1.4	223	2	F85941	probable lipoprote
194	6	1.4	229	2	S64049	hypothetical prote
195	6	1.4	232	2	S42392	G-box-binding prot
196	6	1.4	232	2	A57507	virginiae butanoli
197	6	1.4	233	2	B64436	cobalamin biosynth
198	6	1.4	234	2	S28998	G protein-coupled
199	6	1.4	234	2	S48500	hypothetical prote
200	6	1.4	236	2	S30106	adenylate isopente
201	6	1.4	236	2	F86398	protein F17L21.6 [
202	6	1.4	236	2	S60941	probable membrane
203	6	1.4	237	2	F84013	branched-chain ami
204	6	1.4	240	1	QQECU1	probable transcrip
205	6	1.4	240	2	G85808	hypothetical prote
206	6	1.4	242	2	A34872	muscle-specific re
207	6	1.4	242	2	A42668	myogenic factor MR
208	6	1.4	242	2	S12385	myf-6 protein - hu
209	6	1.4	243	2	S36346	opacity protein op
210	6	1.4	244	2	T50613	hypothetical prote
211	6	1.4	245	2	H71318	conserved hypothet
212	6	1.4	248	2	G81274	probable cyclase C
213	6	1.4	248	2	JC5479	hypothetical prote
214	6	1.4	248	2	F83524	conserved hypothet
215	6	1.4	252	2	T00655	hypothetical prote
216	6	1.4	254	2	B82628	transcription regu
217	6	1.4	255	2	A75507	conserved hypothet
218	6	1.4	255	2	E82813	hypothetical prote
219	6	1.4	257	2	S43584	ribosomal protein
220	6	1.4	258	2	A69830	hypothetical prote
221	6	1.4	259	1	ONGAOL	ovulation hormone
222	6	1.4	260	2	T44854	molybdate-binding
223	6	1.4	260	2	C83686	hypothetical prote
224	6	1.4	264	2	F64736	3-methyl-2-oxobuta
225	6	1.4	264	2	B85497	hypothetical prote
226	6	1.4	264	2	A82833	indole-3-glycerol
227	6	1.4	265	2	D82871	conserved hypothet
228	6	1.4	266	2	S16616	opacity protein op
229	6	1.4	267	2	G75377	conserved hypothet
230	6	1.4	267	2	G84004	spore outer layer
231	6	1.4	268	1	KONH2C	opacity protein P.
232	6	1.4	269	1	OTBY3	cytochrome-c oxida
233	6	1.4	269	2	S78686	cytochrome-c oxida
234	6	1.4	270	2	S04380	opacity protein P.
235	6	1.4	271	2	JQ0950	ICP 18.5 protein -
236	6	1.4	271	2	H72261	conserved hypothet
237	6	1.4	273	2	C64589	hypothetical prote
238	6	1.4	273	2	T40611	hypothetical prote
239	6	1.4	274	2	G69460	conserved hypothet
240	6	1.4	276	2	I53161	alpha 2-adrenergic
241	6	1.4	277	2	A75082	phosphate abc tran
242	6	1.4	277	2	G71456	hypothetical prote
243	6	1.4	278	2	A70034	maltodextrin trans
244	6	1.4	278	2	T44412	ABC transporter (A
245	6	1.4	278	2	E83908	hypothetical prote
246	6	1.4	279	2	F84112	transcription regu
247	6	1.4	281	2	F85956	probable transposa
248	6	1.4	282	2	S16617	opacity protein op

249	6	1.4	284	2	B75386	phosphatidate cyti
250	6	1.4	286	2	A35608	nifC protein - Clo
251	6	1.4	287	2	S71192	mitosis-specific c
252	6	1.4	288	2	D64927	probable shikimate
253	6	1.4	288	2	B43253	zinc finger protei
254	6	1.4	291	1	JN0631	protein kinase (EC
255	6	1.4	292	2	G64695	conserved hypothet
256	6	1.4	292	2	A71824	probable ribonucle
257	6	1.4	292	2	F82385	hypothetical prote
258	6	1.4	293	2	H65182	glucose-1-phosphat
259	6	1.4	293	2	D86065	glucose-1-phosphat
260	6	1.4	293	2	D64900	membrane protein y
261	6	1.4	293	2	D85729	hypothetical prote
262	6	1.4	294	2	G85474	homeodomain-like p
263	6	1.4	295	2	S35914	regulatory protein
264	6	1.4	299	2	T21685	hypothetical prote
265	6	1.4	301	1	E70415	conserved hypothet
266	6	1.4	301	2	T06112	homeotic protein T
267	6	1.4	302	2	G70614	hypothetical prote
268	6	1.4	302	2	A84331	hypothetical prote
269	6	1.4	303	2	H83849	hypothetical prote
270	6	1.4	305	2	S40927	hypothetical prote
271	6	1.4	305	2	F84434	lectin-like protei
272	6	1.4	306	2	B71439	hypothetical prote
273	6	1.4	307	2	C71967	fructose-bisphosph
274	6	1.4	310	2	I64091	cell division prot
275	6	1.4	310	2	E84612	homeodomain transc
276	6	1.4	310	2	C83800	ferrochelatase hem
277	6	1.4	311	1	A64623	thioredoxin reduct
278	6	1.4	311	2	D71890	thioredoxin reduct
279	6	1.4	311	2	S47136	homeotic protein A
280	6	1.4	312	2	F84949	ribose-phosphate p
281	6	1.4	312	2	A46750	olfactory receptor
282	6	1.4	312	2	C86890	competence protein
283	6	1.4	313	2	JC5342	Na+/H+ antiporter
284	6	1.4	314	2	T36913	probable integral
285	6	1.4	315	1	KIEBRT	ribose-phosphate p
286	6	1.4	315	1	KIECRY	ribose-phosphate p
287	6	1.4	315	2	C64132	ribose-phosphate p
288	6	1.4	315	2	E85700	phosphoribosylpyro
289	6	1.4	316	2	T29745	hypothetical prote
290	6	1.4	317	2	C82411	hypothetical prote
291	6	1.4	317	2	T20018	hypothetical prote
292	6	1.4	317	2	T27994	hypothetical prote
293	6	1.4	318	2	S38511	adenosine receptor
294	6	1.4	318	2	E71054	probable sodium de
295	6	1.4	318	2	B85852	probable regulator
296	6	1.4	319	2	JC2252	chitinase (EC 3.2.
297	6	1.4	321	2	T29419	hypothetical prote
298	6	1.4	322	2	T11101	NADH dehydrogenase
299	6	1.4	323	1	QQBED3	HHRF3 protein - hu
300	6	1.4	323	2	T03614	chitinase (EC 3.2.
301	6	1.4	323	2	T20507	hypothetical prote
302	6	1.4	323	2	E75459	probable phosphotr
303	6	1.4	324	2	C70861	probable nrdG prot
304	6	1.4	324	2	T02109	calmodulin-like pr
305	6	1.4	325	2	F86746	hypothetical prote
306	6	1.4	325	2	T23208	hypothetical prote
307	6	1.4	326	2	A59241	DNA-directed DNA p
308	6	1.4	328	2	A82180	chemotaxis protein
309	6	1.4	328	2	G45774	odorant receptor 2
310	6	1.4	328	2	D83275	DNA polymerase III
311	6	1.4	330	2	T50135	probable l-lactate
312	6	1.4	331	2	T36969	probable phytoene
313	6	1.4	331	2	S09800	hypothetical prote
314	6	1.4	332	2	B71508	hypothetical prote
315	6	1.4	333	2	G64188	transmembrane prot
316	6	1.4	335	2	F72096	3-oxoacyl-(acyl-ca
317	6	1.4	335	2	B86528	oxoacyl carrier pr
318	6	1.4	335	2	B36769	11R protein - huma
319	6	1.4	335	2	T50290	calmodulin kinase
320	6	1.4	336	2	S38239	hypothetical prote
321	6	1.4	338	1	S02804	glyceraldehyde-3-p
322	6	1.4	338	2	T41021	ser-thr protein ki
323	6	1.4	338	2	B75154	hypothetical prote
324	6	1.4	338	2	T02645	hypothetical prote
325	6	1.4	340	2	T10516	fruit bromelain (E
326	6	1.4	340	2	I36856	B18L protein - var
327	6	1.4	340	2	T28614	hypothetical prote
328	6	1.4	340	2	B72174	D7L protein - vari
329	6	1.4	340	2	T37451	probable 39.6K pro
330	6	1.4	340	2	G42527	B17L protein - vac
331	6	1.4	341	2	C83067	hypothetical prote
332	6	1.4	341	2	T41450	hypothetical prote
333	6	1.4	342	2	T16806	hypothetical prote
334	6	1.4	343	1	C69211	conserved hypothet
335	6	1.4	344	2	D72116	hypothetical prote
336	6	1.4	344	2	C86507	hypothetical prote
337	6	1.4	344	2	T23028	hypothetical prote
338	6	1.4	345	2	S69018	hypothetical prote
339	6	1.4	346	2	A81554	hypothetical prote
340	6	1.4	347	2	T20103	hypothetical prote
341	6	1.4	348	2	T40017	serine threonine p
342	6	1.4	349	2	S57538	END3 protein - yea
343	6	1.4	350	2	T18850	hypothetical prote
344	6	1.4	352	2	T10501	fruit bromelain (E
345	6	1.4	353	2	A82405	hypothetical prote
346	6	1.4	354	2	T09353	G protein-coupled
347	6	1.4	355	2	B69102	hypothetical prote
348	6	1.4	356	1	S03320	regulatory protein
349	6	1.4	356	2	E71252	probable glycerol-
350	6	1.4	356	2	T19715	hypothetical prote
351	6	1.4	356	2	B86058	probable fimbrial
352	6	1.4	357	2	T37154	hypothetical prote
353	6	1.4	357	2	T21332	hypothetical prote
354	6	1.4	358	2	E81991	NADH dehydrogenase
355	6	1.4	358	2	E81219	NADH dehydrogenase
356	6	1.4	359	2	B26135	keratin, 52K type
357	6	1.4	360	2	T16090	hypothetical prote
358	6	1.4	361	2	T49587	7alpha-cephem-meth
359	6	1.4	362	2	G69589	3-dehydroquinate s
360	6	1.4	362	2	A64212	protein serine/thr
361	6	1.4	364	2	JC5529	p38 beta2 mitogen
362	6	1.4	365	2	T36806	probable alcohol d
363	6	1.4	365	2	C69005	succinyl-CoA synth
364	6	1.4	369	1	S51983	FUN11 protein - ye
365	6	1.4	369	2	C84160	hypothetical prote
366	6	1.4	370	2	T47762	hypothetical prote
367	6	1.4	371	2	T04253	peroxidase homolog
368	6	1.4	372	2	G02524	MAP kinase (EC 2.7
369	6	1.4	376	2	B84994	hypothetical prote
370	6	1.4	376	2	A48840	extradenticle prot
371	6	1.4	377	2	B70140	hypothetical prote
372	6	1.4	378	1	QXBY33	oxi3 intron 3 prot
373	6	1.4	379	2	B86426	hypothetical prote
374	6	1.4	379	2	S55845	3-isopropylmalate
375	6	1.4	381	2	T47555	hypothetical prote
376	6	1.4	383	2	A43470	aliphatic nitrilas
377	6	1.4	383	2	A69065	conserved hypothet
378	6	1.4	384	2	T38544	probable exopolyph
379	6	1.4	385	2	S46529	polygalacturonase
380	6	1.4	385	2	S46530	polygalacturonase
381	6	1.4	385	2	S46532	polygalacturonase
382	6	1.4	385	2	T10003	recF protein - Myc
383	6	1.4	386	2	A70740	probable fadE14 pr
384	6	1.4	386	2	T38150	aminoalcoholphosph
385	6	1.4	388	2	S26824	hypothetical prote
386	6	1.4	390	2	T51201	hypothetical prote
387	6	1.4	393	2	S75810	rod shape-determin
388	6	1.4	394	2	D65167	probable membrane
389	6	1.4	395	2	D72339	hypothetical prote
390	6	1.4	395	2	A86166	protein F21B7.6 [i
391	6	1.4	395	2	T01392	leucine-rich repea
392	6	1.4	396	1	B69851	cytochrome P450 yj
393	6	1.4	396	2	S46533	polygalacturonase
394	6	1.4	397	2	JC4548	argininosuccinate

395	6	1.4	397	2	F64434	hypothetical prote
396	6	1.4	397	2	T00914	leucine-rich repea
397	6	1.4	398	2	A58922	esterase/N-deacety
398	6	1.4	399	2	G75024	hypothetical prote
399	6	1.4	400	2	D86029	probable resistanc
400	6	1.4	402	2	S47768	hypothetical 43K p
401	6	1.4	403	2	JQ0028	cytokeratin 19 - m
402	6	1.4	403	2	B69338	conserved hypothet
403	6	1.4	403	2	T28416	ORF MSV255 leucine
404	6	1.4	404	2	T19831	hypothetical prote
405	6	1.4	406	2	F64074	tyrosine-specific
406	6	1.4	406	2	T52570	1-deoxy-d-xylulose
407	6	1.4	410	1	WMVZF7	BamHI-ORF7 protein
408	6	1.4	411	2	T44250	creatinase (EC 3.5
409	6	1.4	411	2	I39809	creatinase (EC 3.5
410	6	1.4	411	2	C82205	methyl-accepting c
411	6	1.4	414	1	T46280	isocitrate dehydro
412	6	1.4	415	2	S55617	hypothetical prote
413	6	1.4	416	1	G64424	histidine--tRNA li
414	6	1.4	417	2	C32185	Frz CD protein - M
415	6	1.4	419	2	T51715	sigma-like factor
416	6	1.4	428	2	S06266	translation elonga
417	6	1.4	428	2	A86026	uptake of C4-dicar
418	6	1.4	428	2	S47749	C4-dicarboxylate t
419	6	1.4	428	2	B75051	hypothetical prote
420	6	1.4	429	2	S34224	cyclin B - long-ta
421	6	1.4	429	2	T00857	probable receptor-
422	6	1.4	430	2	T28318	ORF MSV157 hypothe
423	6	1.4	432	2	S30433	keratin 17, type I
424	6	1.4	432	2	I50829	alpha 2-adrenocept
425	6	1.4	434	2	D72390	conserved hypothet
426	6	1.4	435	1	S65982	transport protein
427	6	1.4	435	1	WZBE50	gene 50 protein -
428	6	1.4	438	2	B81410	probable integral
429	6	1.4	439	2	H69783	conserved hypothet
430	6	1.4	440	2	S51614	Algal-CAM - Volvox
431	6	1.4	440	2	T51278	hypothetical prote
432	6	1.4	441	2	T01000	hypothetical prote
433	6	1.4	448	2	I51883	alpha-2B-adrenergi
434	6	1.4	450	2	T35873	zinc proteinase -
435	6	1.4	450	2	A38316	alpha-2-adrenergic
436	6	1.4	450	2	B40392	alpha-2-adrenergic
437	6	1.4	450	2	JH0190	alpha-2-adrenergic
438	6	1.4	450	2	A34169	alpha-2A-adrenergi
439	6	1.4	450	2	A37223	alpha-2B-adrenergi
440	6	1.4	450	2	I49481	alpha-2 adrenergic
441	6	1.4	453	2	A35642	alpha-2B-adrenergi
442	6	1.4	455	2	S28221	alpha-2-C2 adrener
443	6	1.4	456	1	ACCHAN	nicotinic acetylch
444	6	1.4	458	2	S55013	NADH dehydrogenase
445	6	1.4	458	2	S67054	probable membrane
446	6	1.4	459	2	T11463	NADH dehydrogenase
447	6	1.4	460	2	T09956	NADH dehydrogenase
448	6	1.4	460	2	S35470	NADH dehydrogenase
449	6	1.4	460	2	A58843	NADH dehydrogenase
450	6	1.4	460	2	B58893	NADH dehydrogenase
451	6	1.4	460	2	T09866	NADH dehydrogenase
452	6	1.4	460	2	T10995	NADH dehydrogenase
453	6	1.4	460	2	B86631	sulfate transporte
454	6	1.4	460	2	I57546	Rabin3 - rat
455	6	1.4	462	2	C85651	hypothetical prote
456	6	1.4	463	1	S75661	membrane-bound pro
457	6	1.4	464	2	T15161	hypothetical prote
458	6	1.4	465	2	B55748	protein kinase (EC
459	6	1.4	465	2	T18793	hypothetical prote
460	6	1.4	467	2	I50476	keratin type I - g
461	6	1.4	467	2	D86237	protein F14N23.21
462	6	1.4	468	2	T12725	terminase large ch
463	6	1.4	470	2	C70554	hypothetical prote
464	6	1.4	471	2	A54739	indole-3-acetate b
465	6	1.4	471	2	B83127	probable secretion
466	6	1.4	472	1	KRHUE	keratin 14, type I
467	6	1.4	472	2	S31459	Ig gamma-1 chain -
468	6	1.4	472	2	C75064	hypothetical prote
469	6	1.4	473	2	E84471	probable beta-1,3-
470	6	1.4	477	2	T15604	hypothetical prote
471	6	1.4	477	2	D83617	probable amino aci
472	6	1.4	478	2	S51457	probable membrane
473	6	1.4	478	4	TVHUBD	transforming prote
474	6	1.4	479	2	F82328	aspartokinase III,
475	6	1.4	479	2	T49871	peroxisomal Ca-dep
476	6	1.4	481	2	C71079	hypothetical prote
477	6	1.4	481	2	T48516	probable oligopept
478	6	1.4	482	1	JN0506	argininosuccinate
479	6	1.4	483	2	E86881	glutamate--tRNA li
480	6	1.4	483	2	T20895	hypothetical prote
481	6	1.4	484	2	T07782	cytochrome-c oxida
482	6	1.4	484	2	E64085	X-His dipeptidase
483	6	1.4	486	1	S64534	hypothetical prote
484	6	1.4	487	1	AJSMRC	argininosuccinate
485	6	1.4	487	2	C82137	FAD monooxygenase,
486	6	1.4	488	2	F84493	hypothetical prote
487	6	1.4	488	2	D86698	SOS responce UmuC
488	6	1.4	489	2	T10283	ecdysteroid UDPglu
489	6	1.4	490	2	B28516	cytochrome P450 2C
490	6	1.4	490	2	T19881	hypothetical prote
491	6	1.4	490	2	F71487	probable DNA gyras
492	6	1.4	490	2	C81748	DNA gyrase, chain
493	6	1.4	490	2	F84154	Na+/H+ antiporter
494	6	1.4	491	2	B96633	hypothetical prote
495	6	1.4	496	2	D75293	uracil permease -
496	6	1.4	496	2	T47884	pectinesterase-lik
497	6	1.4	498	2	E69033	NADH dehydrogenase
498	6	1.4	498	2	H83880	hypothetical prote
499	6	1.4	499	2	T05893	lysosomal Pro-X ca
500	6	1.4	500	1	S37284	cytochrome P450 2D
501	6	1.4	500	2	G84706	hypothetical prote
502	6	1.4	501	2	I61512	TNF receptor assoc
503	6	1.4	504	2	S70986	replication initia
504	6	1.4	504	2	I56542	calmodulin-binding
505	6	1.4	505	2	T01667	gag polyprotein -
506	6	1.4	506	2	H86346	hypothetical prote
507	6	1.4	507	2	B42249	serine-type carbox
508	6	1.4	509	1	S22339	steroid 17alpha-mo
509	6	1.4	509	2	T38890	pyruvate kinase (E
510	6	1.4	509	2	T45166	pyruvate kinase (E
511	6	1.4	513	1	APBOL	leucyl aminopeptid
512	6	1.4	513	2	T21887	hypothetical prote
513	6	1.4	514	2	T04155	glucose-1-phosphat
514	6	1.4	515	2	C75615	exopolyphosphatase
515	6	1.4	516	2	T40631	probable membrane
516	6	1.4	517	2	T23413	hypothetical prote
517	6	1.4	518	2	B71545	probable oligopept
518	6	1.4	519	2	T39822	aspartate kinase -
519	6	1.4	519	2	C71346	probable transcrip
520	6	1.4	520	2	T16593	hypothetical prote
521	6	1.4	522	2	T27993	hypothetical prote
522	6	1.4	523	1	NIAVMB	nitrogenase (EC 1.
523	6	1.4	523	2	T15649	hypothetical prote
524	6	1.4	525	2	G71317	hypothetical prote
525	6	1.4	526	2	F84385	dipeptide ABC tran
526	6	1.4	530	2	T43089	transfer complex p
527	6	1.4	531	2	T12406	cytochrome-c oxida
528	6	1.4	531	2	I39692	cytochrome-c3 hydr
529	6	1.4	531	2	G83268	conserved hypothet
530	6	1.4	531	2	T04463	hypothetical prote
531	6	1.4	534	1	A48529	mitochondrial proc
532	6	1.4	534	2	I37942	keratin 4, type II
533	6	1.4	534	2	S31300	regulatory protein
534	6	1.4	534	2	JC5096	transposase - fung
535	6	1.4	536	2	T07946	cytochrome-c oxida
536	6	1.4	536	2	A81710	ribosomal protein
537	6	1.4	538	2	G84129	two-component sens
538	6	1.4	538	2	B72609	hypothetical prote
539	6	1.4	540	2	T34702	hypothetical prote
540	6	1.4	540	2	T25851	hypothetical prote

541	6	1.4	543	1	JC4070	protein kinase (EC
542	6	1.4	543	2	S29355	Photinus-luciferin
543	6	1.4	543	2	S29352	Photinus-luciferin
544	6	1.4	543	2	S29354	Photinus-luciferin
545	6	1.4	543	2	S29353	Photinus-luciferin
546	6	1.4	543	2	A81254	probable ABC trans
547	6	1.4	543	2	T20964	hypothetical prote
548	6	1.4	545	2	F86834	alpha-glucosidase
549	6	1.4	545	2	T09195	secY protein homol
550	6	1.4	546	1	COBYC1	cyclin 1 - yeast (
551	6	1.4	546	2	D83408	hypothetical prote
552	6	1.4	547	2	T22856	hypothetical prote
553	6	1.4	552	2	G69464	acetolactate synth
554	6	1.4	554	2	S48493	malate synthase (E
555	6	1.4	560	1	HMIVT7	hemagglutinin prec
556	6	1.4	562	2	C71604	RNA helicase PFB08
557	6	1.4	566	2	S62482	serine/threonine p
558	6	1.4	567	2	S10138	mcf2 protein - hum
559	6	1.4	567	2	D84029	two-component sens
560	6	1.4	569	2	A71556	probable S1 riboso
561	6	1.4	569	2	T48575	aspartate kinase -
562	6	1.4	571	1	UMFF	period clock prote
563	6	1.4	571	2	E64903	arylsulfatase homo
564	6	1.4	571	2	B85726	probable sulfatase
565	6	1.4	574	1	S11032	5'-nucleotidase (E
566	6	1.4	580	2	C86530	S1 ribosomal prote
567	6	1.4	580	2	D72093	ribosomal protein
568	6	1.4	587	2	S33951	fiber protein - hu
569	6	1.4	588	2	E96633	probable Serine/Th
570	6	1.4	590	2	E72015	DNA primase - Chla
571	6	1.4	590	2	D86608	DNA primase [impor
572	6	1.4	590	2	B84898	hypothetical prote
573	6	1.4	591	2	A69291	transmembrane olig
574	6	1.4	594	2	A81525	DNA primase, proba
575	6	1.4	594	2	T42660	hypothetical prote
576	6	1.4	594	2	I49127	intracellular prot
577	6	1.4	597	2	F82543	outer membrane hem
578	6	1.4	598	2	S51386	hypothetical prote
579	6	1.4	601	2	T34396	hypothetical prote
580	6	1.4	601	2	C81715	gcpE protein TC032
581	6	1.4	602	2	T11438	NADH dehydrogenase
582	6	1.4	603	1	QQBED1	HHLF5 protein - hu
583	6	1.4	603	2	T11284	NADH dehydrogenase
584	6	1.4	606	2	E84889	hypothetical prote
585	6	1.4	609	2	T40660	hypothetical C2H2-
586	6	1.4	609	2	S64539	probable membrane
587	6	1.4	611	2	T44880	probable acyl-coA
588	6	1.4	612	2	S35471	NADH dehydrogenase
589	6	1.4	612	2	T13616	hypothetical prote
590	6	1.4	613	2	S27770	hypothetical prote
591	6	1.4	616	2	G64160	GTP-binding protei
592	6	1.4	616	2	C75588	conserved hypothet
593	6	1.4	623	2	T39001	probable exonuclea
594	6	1.4	624	2	A83237	hypothetical prote
595	6	1.4	624	2	A84150	hypothetical prote
596	6	1.4	627	2	S68958	carnitine O-acetyl
597	6	1.4	628	2	S37795	hypothetical prote
598	6	1.4	630	2	S49969	probable membrane
599	6	1.4	631	1	BHTLA	hemocyanin chain a
600	6	1.4	633	2	F84865	hypothetical prote
601	6	1.4	633	2	T05005	hypothetical prote
602	6	1.4	634	2	I40217	glutaryl 7-ACA acy
603	6	1.4	635	1	A64162	cytochrome c-type
604	6	1.4	638	2	T38863	hypothetical prote
605	6	1.4	640	2	C72351	hypothetical prote
606	6	1.4	641	2	G28771	hypothetical prote
607	6	1.4	641	2	C82206	methyl-accepting c
608	6	1.4	644	2	S44478	NADH dehydrogenase
609	6	1.4	646	2	H82555	c-type cytochrome
610	6	1.4	647	2	JQ0033	site-specific DNA-
611	6	1.4	649	2	T24505	hypothetical prote
612	6	1.4	649	2	T38883	hypothetical prote
613	6	1.4	650	2	T48060	beta-D-glucan exoh
614	6	1.4	654	1	BVBYP1	CBP1 protein - yea
615	6	1.4	655	2	A48468	dnaK-type molecula
616	6	1.4	657	1	BVECGX	glycogen operon pr
617	6	1.4	657	2	E86009	hypothetical prote
618	6	1.4	658	2	T16040	hypothetical prote
619	6	1.4	663	2	A86626	hypothetical prote
620	6	1.4	663	2	D83545	probable helicase
621	6	1.4	664	2	S44756	probable protein d
622	6	1.4	664	2	G86663	hypothetical prote
623	6	1.4	666	1	JE0201	protein-tyrosine k
624	6	1.4	667	2	JC5889	OS-9 protein precu
625	6	1.4	671	2	H81976	probable prolyl ol
626	6	1.4	671	2	C81033	prolyl oligopeptid
627	6	1.4	675	2	S53831	NADH dehydrogenase
628	6	1.4	676	1	A39379	hatching-suppresse
629	6	1.4	678	1	RDRTO4	NADPH--ferrihemopr
630	6	1.4	678	2	T02867	peptidyl dipeptida
631	6	1.4	679	2	G83492	hypothetical prote
632	6	1.4	682	2	F64704	DNA helicase II -
633	6	1.4	682	2	S71476	homeotic protein H
634	6	1.4	683	2	B86354	protein F2E2.22 [i
635	6	1.4	683	2	D71433	hypothetical prote
636	6	1.4	684	2	H96646	hypothetical prote
637	6	1.4	684	2	T02149	hypothetical prote
638	6	1.4	687	2	S43587	F26F3.2 protein (c
639	6	1.4	691	2	C72454	hypothetical prote
640	6	1.4	699	2	B72775	probable DNA repli
641	6	1.4	703	2	T05632	hypothetical prote
642	6	1.4	705	2	T24343	hypothetical prote
643	6	1.4	707	1	A53530	protein kinase C (
644	6	1.4	708	2	E84950	phosphate acetyltr
645	6	1.4	710	2	S66313	phenylalanine ammo
646	6	1.4	720	2	JC6562	heat shock protein
647	6	1.4	721	2	T09631	probable acylamino
648	6	1.4	727	1	A38206	procollagen-lysine
649	6	1.4	727	2	T17292	hypothetical prote
650	6	1.4	728	2	S59964	procollagen-lysine
651	6	1.4	728	2	T51071	related to trfA pr
652	6	1.4	730	2	A23742	procollagen-lysine
653	6	1.4	732	2	T47269	copper-transportin
654	6	1.4	732	2	T49087	hypothetical prote
655	6	1.4	735	2	T35778	probable fusidic a
656	6	1.4	739	2	I40715	malate synthase (E
657	6	1.4	740	2	A75011	hypothetical prote
658	6	1.4	743	2	S48917	probable regulator
659	6	1.4	743	2	D84854	hypothetical prote
660	6	1.4	748	2	C64409	hypothetical prote
661	6	1.4	755	2	S32103	filensin - bovine
662	6	1.4	757	2	B72273	hypothetical prote
663	6	1.4	760	2	T01441	hypothetical prote
664	6	1.4	761	2	T34038	hypothetical prote
665	6	1.4	763	2	H70145	ATP-dependent Clp
666	6	1.4	764	2	C83513	hypothetical prote
667	6	1.4	765	2	G64502	hypothetical prote
668	6	1.4	765	2	G85039	hypothetical prote
669	6	1.4	766	2	F83102	probable ferrous i
670	6	1.4	766	2	JC5958	sperm surface prot
671	6	1.4	768	2	D84271	small heat shock p
672	6	1.4	768	2	F75350	probable protein-e
673	6	1.4	769	2	F81742	conserved hypothet
674	6	1.4	770	1	WZBE30	gene 30 protein -
675	6	1.4	770	2	H85738	hypothetical prote
676	6	1.4	771	2	T16254	hypothetical prote
677	6	1.4	778	2	H96649	protein F24O1.4 [i
678	6	1.4	779	2	T31732	hypothetical prote
679	6	1.4	783	2	E86254	hypothetical prote
680	6	1.4	784	1	SUECLA	endopeptidase La (
681	6	1.4	784	1	S47270	endopeptidase La (
682	6	1.4	786	2	G82141	ATP-dependent LA p
683	6	1.4	786	2	A69308	DNA gyrase, subuni
684	6	1.4	788	2	S05661	muscarinic acetylc
685	6	1.4	789	2	T51310	RNA helicase RH28
686	6	1.4	793	1	JC6161	kinesin-associated

687	6	1.4	795	2	E86247	receptor-like prot
688	6	1.4	796	2	S50672	hypothetical prote
689	6	1.4	799	2	A85541	hypothetical prote
690	6	1.4	800	2	S65046	1,4-alpha-glucan b
691	6	1.4	804	2	A37473	probable RNA polym
692	6	1.4	805	2	T48392	1,4-alpha-glucan b
693	6	1.4	810	2	D64090	glycerol-3-phospha
694	6	1.4	810	2	T33323	hypothetical prote
695	6	1.4	811	2	B69512	cell division cont
696	6	1.4	811	2	F83451	probable cation-tr
697	6	1.4	811	2	T25166	hypothetical prote
698	6	1.4	813	2	JC5809	hypoxia-inducible
699	6	1.4	815	2	T05555	DNA polymerase III
700	6	1.4	821	2	T24728	hypothetical prote
701	6	1.4	825	2	A26983	regulatory protein
702	6	1.4	826	2	I38972	hypoxia-inducible
703	6	1.4	827	2	S29955	surface glycoprote
704	6	1.4	829	2	T46556	DNA topoisomerase
705	6	1.4	831	2	S44843	K06H7.3 protein -
706	6	1.4	832	2	G86147	protein T1N6.6 [im
707	6	1.4	833	2	H81700	DNA gyrase, chain
708	6	1.4	833	2	T32289	hypothetical prote
709	6	1.4	834	2	B86525	DNA gyrase subunit
710	6	1.4	834	2	G72098	DNA gyrase chain A
711	6	1.4	834	2	F81571	DNA gyrase, chain
712	6	1.4	836	2	F71546	probable DNA gyras
713	6	1.4	836	2	T46070	hypothetical prote
714	6	1.4	837	2	S54624	ROD1 protein - yea
715	6	1.4	838	2	I45557	eyeless, long form
716	6	1.4	846	2	T38840	hypothetical trp-a
717	6	1.4	848	2	C64841	trimethylamine-N-o
718	6	1.4	848	2	D85635	trimethylamine N-o
719	6	1.4	852	2	B72685	hypothetical prote
720	6	1.4	855	1	JQ2004	env polyprotein -
721	6	1.4	857	2	E71377	probable ATP-depen
722	6	1.4	858	2	S30571	DNA topoisomerase
723	6	1.4	869	2	G83422	aconitate hydratas
724	6	1.4	873	2	T19131	hypothetical prote
725	6	1.4	874	2	T30398	late expression fa
726	6	1.4	877	2	T43449	hypothetical prote
727	6	1.4	896	2	A41273	glutamate receptor
728	6	1.4	906	2	A40222	glutamate receptor
729	6	1.4	906	2	S38723	glutamate receptor
730	6	1.4	906	2	S25852	glutamate receptor
731	6	1.4	907	1	ACRTK1	glutamate receptor
732	6	1.4	907	2	A40170	glutamate receptor
733	6	1.4	907	2	S12874	glutamate receptor
734	6	1.4	911	2	S51441	hypothetical prote
735	6	1.4	916	2	T03323	gene 116 protein -
736	6	1.4	920	2	T43263	cell division prot
737	6	1.4	925	1	TVHUDB	transforming prote
738	6	1.4	925	2	T37475	lipoprotein recept
739	6	1.4	925	2	T07713	probable ABC-type
740	6	1.4	926	2	S36742	cation-transportin
741	6	1.4	942	2	C96574	hypothetical prote
742	6	1.4	956	2	S30834	hypothetical prote
743	6	1.4	965	2	T09846	phosphoenolpyruvat
744	6	1.4	966	2	G69189	hypothetical prote
745	6	1.4	968	2	B64385	malT-glpR intergen
746	6	1.4	976	2	S40697	processing endopro
747	6	1.4	976	2	T29583	hypothetical prote
748	6	1.4	977	2	C96745	hypothetical prote
749	6	1.4	1009	2	T45645	receptor kinase-li
750	6	1.4	1012	1	GNXSOH	genome polyprotein
751	6	1.4	1013	2	S32214	hypothetical prote
752	6	1.4	1021	2	T10748	mannan endo-1,4-be
753	6	1.4	1029	2	S56229	probable membrane
754	6	1.4	1042	2	H70203	isoleucine--tRNA l
755	6	1.4	1042	2	T26644	hypothetical prote
756	6	1.4	1047	2	S19508	MSH3 protein - yea
757	6	1.4	1054	2	T30933	chitinase (EC 3.2.
758	6	1.4	1063	2	A40253	acidic nuclear pro
759	6	1.4	1065	2	T25068	hypothetical prote
760	6	1.4	1073	1	A25526	ring-infected eryt
761	6	1.4	1073	2	T33764	hypothetical prote
762	6	1.4	1082	2	S37837	suppressor protein
763	6	1.4	1092	2	B86401	protein T22C5.3 [i
764	6	1.4	1099	2	A56155	tumor suppressor p
765	6	1.4	1110	2	A43253	Large tra-1 protei
766	6	1.4	1112	2	D75056	cell division cont
767	6	1.4	1122	2	B26427	period clock prote
768	6	1.4	1126	2	T20801	hypothetical prote
769	6	1.4	1127	2	A25018	circadian rhythm p
770	6	1.4	1127	2	G71274	hypothetical prote
771	6	1.4	1133	1	GNVUSR	M polyprotein prec
772	6	1.4	1133	1	A43964	M polyprotein prec
773	6	1.4	1133	2	S12597	M polyprotein prec
774	6	1.4	1134	1	GNVU22	M polyprotein prec
775	6	1.4	1134	1	A43960	M polyprotein prec
776	6	1.4	1134	2	T04587	hypothetical prote
777	6	1.4	1140	1	I38908	UV-damaged DNA-bin
778	6	1.4	1140	1	S38777	UV-damaged DNA-bin
779	6	1.4	1140	2	JC7152	UV-damaged DNA-bin
780	6	1.4	1167	2	T06146	disease resistance
781	6	1.4	1175	2	T46124	hypothetical prote
782	6	1.4	1176	2	C26427	period clock prote
783	6	1.4	1178	2	S57535	probable membrane
784	6	1.4	1190	2	S47536	SWH1 protein - yea
785	6	1.4	1192	2	T48499	receptor-like prot
786	6	1.4	1206	2	T18557	probable hydrogena
787	6	1.4	1208	2	S17286	period clock prote
788	6	1.4	1215	2	E70614	hypothetical prote
789	6	1.4	1218	2	A26427	period clock prote
790	6	1.4	1218	2	A26588	period clock prote
791	6	1.4	1221	2	T25005	hypothetical prote
792	6	1.4	1226	2	T45788	probable disease r
793	6	1.4	1234	2	T14515	hypothetical prote
794	6	1.4	1248	2	A47445	reverse gyrase - S
795	6	1.4	1251	2	B86194	hypothetical prote
796	6	1.4	1263	2	T13805	spalt-related prot
797	6	1.4	1268	2	A49674	flightless-I homol
798	6	1.4	1275	2	JU0092	trp protein - frui
799	6	1.4	1279	1	E69681	peptide synthetase
800	6	1.4	1344	2	T14316	rig-1 protein - mo
801	6	1.4	1356	2	E86490	hypothetical prote
802	6	1.4	1389	2	T03273	embryogenesis tran
803	6	1.4	1411	2	S30355	alpha-latroinsecto
804	6	1.4	1414	2	T33236	hypothetical prote
805	6	1.4	1415	2	C83070	conserved hypothet
806	6	1.4	1447	2	F82909	hypothetical prote
807	6	1.4	1451	2	S41025	hypothetical prote
808	6	1.4	1507	2	T42631	breast cancer tumo
809	6	1.4	1513	2	A70982	probable ATP-depen
810	6	1.4	1541	1	S71839	canalicular multid
811	6	1.4	1545	1	S71841	multidrug resistan
812	6	1.4	1553	2	T09361	hypothetical prote
813	6	1.4	1608	2	T13216	minor capsid prote
814	6	1.4	1613	2	E82193	conserved hypothet
815	6	1.4	1642	2	I51018	cobra venom factor
816	6	1.4	1717	2	T50247	probable helicase
817	6	1.4	1825	2	T32828	hypothetical prote
818	6	1.4	1839	1	RRWPEM	genome polyprotein
819	6	1.4	1872	2	S64851	probable membrane
820	6	1.4	1927	2	T25604	hypothetical prote
821	6	1.4	1940	2	F75393	hypothetical prote
822	6	1.4	1979	2	C71622	hypothetical prote
823	6	1.4	1984	2	A44396	P-type cation tran
824	6	1.4	1998	2	T13009	hypothetical prote
825	6	1.4	2032	2	I39917	hypothetical prote
826	6	1.4	2055	2	T00093	hypothetical prote
827	6	1.4	2139	2	A44467	voltage-dependent
828	6	1.4	2143	2	JH0427	voltage-dependent
829	6	1.4	2161	1	A45389	genome polyprotein
830	6	1.4	2183	1	ZLNZMV	genome polyprotein
831	6	1.4	2183	1	G48556	genome polyprotein
832	6	1.4	2183	2	S47307	genome polyprotein

833	6	1.4	2201	1	GNNYA9	genome polyprotein
834	6	1.4	2205	2	T08615	aggregation factor
835	6	1.4	2210	1	RRXPLC	genome polyprotein
836	6	1.4	2236	1	QZFF	rudimentary protei
837	6	1.4	2249	2	A41477	190K surface antig
838	6	1.4	2287	2	T21312	hypothetical prote
839	6	1.4	2297	2	T34918	polyketide synthas
840	6	1.4	2298	2	T49648	hypothetical prote
841	6	1.4	2471	2	T03820	probable histidine
842	6	1.4	2672	2	A48126	translation activa
843	6	1.4	2895	2	T08437	hyperplastic discs
844	6	1.4	3034	2	T14119	seven-pass transme
845	6	1.4	3144	2	S64791	VPS13 protein - ye
846	6	1.4	3429	2	T13853	hypothetical prote
847	6	1.4	4085	2	S28600	hypothetical prote
848	6	1.4	4180	2	G83559	hypothetical prote
849	6	1.4	4466	1	S17231	dynein beta heavy
850	6	1.4	4466	1	S17653	dynein beta heavy
851	5	1.2	14	2	PH1305	Ig heavy chain DJ
852	5	1.2	14	2	A43847	Ig mu chain V regi
853	5	1.2	15	2	PS0221	gastrin-releasing
854	5	1.2	15	2	B26501	lipoprotein lipase
855	5	1.2	16	2	E53284	T-cell receptor be
856	5	1.2	19	2	I53673	amyloid protein -
857	5	1.2	21	2	S16073	alanine--tRNA liga
858	5	1.2	23	2	PC4030	rRNA endonuclease
859	5	1.2	24	2	A37825	fibronectin recept
860	5	1.2	25	2	T01689	ATPase-beta chain
861	5	1.2	27	2	H48200	homeotic protein H
862	5	1.2	30	2	C42842	antifungal 2S stor
863	5	1.2	30	2	I37626	Fc gamma (IgG) rec
864	5	1.2	31	2	T08489	hypothetical prote
865	5	1.2	31	2	T11991	hypothetical prote
866	5	1.2	32	2	S20771	Ig heavy chain V r
867	5	1.2	33	2	PC4352	cytochrome P450 4A
868	5	1.2	34	2	F70242	hypothetical prote
869	5	1.2	35	2	G70139	hypothetical prote
870	5	1.2	35	2	D82224	hypothetical prote
871	5	1.2	35	2	D82607	hypothetical prote
872	5	1.2	37	2	S16183	napin nIII - rape
873	5	1.2	37	2	S35591	major allergen sma
874	5	1.2	38	1	F2RZL	photosystem II pro
875	5	1.2	38	1	F2WTL	photosystem II pro
876	5	1.2	38	1	F2SKL	photosystem II pro
877	5	1.2	38	1	F2KTL	photosystem II pro
878	5	1.2	38	2	T07254	photosystem II pro
879	5	1.2	38	2	S78339	photosystem II pro
880	5	1.2	38	2	S04064	photosystem II pro
881	5	1.2	38	2	S05685	photosystem II pro
882	5	1.2	38	2	S31821	photosystem II pro
883	5	1.2	38	2	A05048	photosystem II pro
884	5	1.2	38	2	S03193	photosystem II pro
885	5	1.2	38	2	S00691	photosystem II pro
886	5	1.2	38	2	S58566	photosystem II pro
887	5	1.2	38	2	C48310	photosystem II pro
888	5	1.2	38	2	S51366	photosystem II pro
889	5	1.2	38	2	S28056	photosystem II pro
890	5	1.2	38	2	S73310	photosystem II pro
891	5	1.2	38	2	T07477	photosystem II pro
892	5	1.2	39	1	C69677	phosphatase (RapF)
893	5	1.2	39	2	PQ0801	calmodulin antagon
894	5	1.2	39	2	PQ0802	calmodulin antagon
895	5	1.2	39	2	PQ0799	calmodulin antagon
896	5	1.2	39	2	JQ2321	calmodulin antagon
897	5	1.2	39	2	S01791	allergen Sin a I s
898	5	1.2	39	2	A38623	collagen alpha 1(I
899	5	1.2	39	2	I52552	connexin 32 - rat
900	5	1.2	39	2	I37554	homeobox - human (
901	5	1.2	40	4	A58361	protein-tyrosine k
902	5	1.2	43	2	B48359	hydroxymethylbilan
903	5	1.2	43	2	JQ1414	hypothetical 4.3K
904	5	1.2	44	2	S26879	photosystem II pro
905	5	1.2	46	2	A48542	CRF-related diuret
906	5	1.2	46	2	F81142	hypothetical prote
907	5	1.2	47	2	G70191	hypothetical prote
908	5	1.2	47	2	F81118	hypothetical prote
909	5	1.2	48	2	H64544	ribosomal protein
910	5	1.2	48	2	G71964	ribosomal protein
911	5	1.2	49	1	Q9BPE7	gene 19.5 protein
912	5	1.2	50	2	S67520	ferredoxin [3Fe-4S
913	5	1.2	50	2	B86576	hypothetical prote
914	5	1.2	50	2	PN0101	hypothetical prote
915	5	1.2	50	2	C72047	hypothetical prote
916	5	1.2	51	1	XKPOT	proteinase inhibit
917	5	1.2	52	1	BVECMA	stable plasmid inh
918	5	1.2	52	2	C38230	inorganic pyrophos
919	5	1.2	52	2	S00301	hok protein - Esch
920	5	1.2	52	2	S33146	H+-transporting AT
921	5	1.2	52	2	S52069	p36 protein (CAK c
922	5	1.2	52	2	E82702	hypothetical prote
923	5	1.2	52	4	T11818	hypothetical Pdlec
924	5	1.2	53	2	S15953	repA protein - Lac
925	5	1.2	53	2	JQ1198	hypothetical 5.8K
926	5	1.2	53	2	T03342	gene e25 protein -
927	5	1.2	53	2	A82862	hypothetical prote
928	5	1.2	54	2	S36013	H+-transporting AT
929	5	1.2	54	2	S35628	myosin heavy chain
930	5	1.2	54	2	C39443	hypothetical prote
931	5	1.2	54	2	T42318	hypothetical prote
932	5	1.2	55	2	JC6188	csgAB protein - Ba
933	5	1.2	56	2	S36380	Ig heavy chain V r
934	5	1.2	56	2	S77849	probable isoleucin
935	5	1.2	56	2	H82667	hypothetical prote
936	5	1.2	57	1	R5EC32	ribosomal protein
937	5	1.2	57	1	W9BPC7	gene 19.3 protein
938	5	1.2	57	2	G85671	50S ribosomal subu
939	5	1.2	58	2	S76479	hypothetical prote
940	5	1.2	58	2	D70502	hypothetical prote
941	5	1.2	58	2	T47817	hypothetical prote
942	5	1.2	58	2	S44902	ZK652.2 protein -
943	5	1.2	58	2	S33413	NuMA protein - hum
944	5	1.2	59	1	A55807	kaliotoxin 2 precu
945	5	1.2	59	2	T43106	hypothetical prote
946	5	1.2	59	2	I40012	sspF protein - Bac
947	5	1.2	60	2	E85790	hypothetical prote
948	5	1.2	60	2	D64940	hypothetical prote
949	5	1.2	61	2	A36853	A45R protein - var
950	5	1.2	61	2	A64005	hypothetical prote
951	5	1.2	61	2	A48175	hypothetical prote
952	5	1.2	61	2	E72169	A48R protein - var
953	5	1.2	61	2	F83792	hypothetical prote
954	5	1.2	62	1	CKWKB	cecropin B precurs
955	5	1.2	62	2	S53002	mitotic-specific c
956	5	1.2	62	2	I39622	ribosomal protein
957	5	1.2	62	2	S73237	hypothetical prote
958	5	1.2	62	2	S53981	hypothetical prote
959	5	1.2	62	2	T51850	RING-H2 finger pro
960	5	1.2	63	1	Q3ECD7	ydfD protein - Esc
961	5	1.2	63	2	S65918	homeotic protein e
962	5	1.2	63	2	G96531	hypothetical prote
963	5	1.2	63	2	S72911	probable phospholi
964	5	1.2	63	2	A69035	hypothetical prote
965	5	1.2	64	2	A41929	cAMP-dependent pro
966	5	1.2	64	2	S20895	homeotic protein c
967	5	1.2	64	2	G71830	hypothetical prote
968	5	1.2	65	2	I46680	apolipoprotein B-4
969	5	1.2	65	2	D42507	F8L protein - vacc
970	5	1.2	65	2	B36213	F8L protein - vacc
971	5	1.2	65	2	T30785	hypothetical prote
972	5	1.2	65	2	C36840	C12L protein - var
973	5	1.2	65	2	F72154	E8L protein - vari
974	5	1.2	66	1	R5BS29	ribosomal protein
975	5	1.2	66	1	R5BS2L	ribosomal protein
976	5	1.2	66	2	S31652	enterotoxin - Yers
977	5	1.2	66	2	BB6648	hypothetical prote
978	5	1.2	66	2	G85910	unknown protein en

979	5	1.2	66	2	A81145	hypothetical prote
980	5	1.2	66	2	C82315	transcription regu
981	5	1.2	67	2	S26155	H+-transporting AT
982	5	1.2	67	2	B72248	ribosomal protein
983	5	1.2	67	2	B81328	hypothetical prote
984	5	1.2	67	2	T16814	FMRFamide-like pep
985	5	1.2	67	2	F83667	ribosomal protein
986	5	1.2	68	2	S12897	serine proteinase
987	5	1.2	68	2	T18175	hypothetical prote
988	5	1.2	68	2	T36853	hypothetical prote
989	5	1.2	68	2	T39638	probable metal hom
990	5	1.2	70	2	S48885	DNA-directed RNA p
991	5	1.2	70	2	C86462	unknown protein, 6
992	5	1.2	70	2	B84300	hypothetical prote
993	5	1.2	70	2	E71878	GTP cyclohydrolase
994	5	1.2	70	2	C81126	hypothetical prote
995	5	1.2	71	2	G81889	hypothetical prote
996	5	1.2	71	2	D81003	hypothetical prote
997	5	1.2	71	2	D82025	probable integral
998	5	1.2	71	2	S42116	hypothetical prote
999	5	1.2	71	2	T33214	hypothetical prote
1000	5	1.2	71	2	I70114	arrestin-D - rat (

ALIGNMENTS

RESULT 1
A48349
UL28 protein - ateline herpesvirus 1 (strain Lennette) (fragment)
C;Species: ateline herpesvirus 1
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 26-Aug-1999
C;Accession: A48349
R;Eberle, R.; Black, D.
Arch. Virol. 129, 167-182, 1993
A;Title: Sequence analysis of herpes simplex virus gB gene homologs of two platyrrhine m
A;Reference number: A48349; MUID:93228440
A;Accession: A48349
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-266 <EBE>
A;Cross-references: GB:M95785; NID:g305364; PIDN:AAA43838.1; PID:g305365
A;Note: sequence extracted from NCBI backbone (NCBIN:129063, NCBIP:129064)
C;Superfamily: herpesvirus infected cell protein ICP18.5

Query Match 1.9%; Score 8; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 LKEELARF 114
||||||||
Db 116 LKEELARF 123

RESULT 2
WMBEBH
infected cell protein ICP18.5 - bovine herpesvirus 2 (strain BMV)
C;Species: bovine herpesvirus 2
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 11-Nov-1996
C;Accession: B29242
R;Hammerschmidt, W.; Conraths, F.; Mankertz, J.; Pauli, G.; Ludwig, H.; Buhk, H.J.
Virology 165, 388-405, 1988
A;Title: Conservation of a gene cluster including glycoprotein B in bovine herpesvirus t
A;Reference number: A94381; MUID:88306231
A;Accession: B29242
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-664 <HAM>
A;Cross-references: GB:M21628
C;Superfamily: herpesvirus infected cell protein ICP18.5
C;Keywords: capsid assembly

Query Match 1.9%; Score 8; DB 1; Length 664;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 LKEELARF 114
||||||||
Db 524 LKEELARF 631

RESULT 3
WMBEPR
infected cell protein ICP18.5 - suid herpesvirus 1 (strain Becker)
C;Species: suid herpesvirus 1
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C;Accession: A33779
R;Pederson, N.E.; Enquist, L.W.
Nucleic Acids Res. 17, 3597, 1989
A;Title: The nucleotide sequence of a pseudorabies virus gene similar to ICP18.5 of h
A;Reference number: S04145; MUID:89263808
A;Accession: A33779
A;Molecule type: DNA
A;Residues: 1-724 <PED>
A;Cross-references: GB:X14573; NID:g61407; PIDN:CAA32712.1; PID:g61409
C;Superfamily: herpesvirus infected cell protein ICP18.5
C;Keywords: capsid assembly; late protein

Query Match 1.9%; Score 8; DB 1; Length 724;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 LKEELARF 114
||||||||
Db 602 LKEELARF 609

RESULT 4
WZBEC5
infected cell protein ICP18.5 - equine herpesvirus 1
C;Species: equine herpesvirus 1
A;Note: host Equus caballus (domestic horse)
C;Date: 30-Sep-1992 #sequence_revision 30-Jun-1993 #text_change 11-Nov-1996
C;Accession: F36798; JQ0847
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
submitted to GenBank, March 1992
A;Description: The DNA sequence of equine herpesvirus-1.
A;Reference number: A36805
A;Accession: F36798
A;Molecule type: DNA
A;Residues: 1-766 <TEL>
A;Cross-references: GB:M86664
A;Experimental source: strain Ab4p
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A;Title: The DNA sequence of equine herpesvirus-1.
A;Reference number: A41831; MUID:92295566
A;Contents: annotation; possible protein-coding frames
A;Note: neither amino acid nor nucleotide sequence is given
R;Bell, C.W.; Whalley, J.M.
submitted to JIPID, January 1991
A;Reference number: JQ0846
A;Accession: JQ0847
A;Molecule type: DNA
A;Residues: 1-766 <BEL>
C;Genetics:
A;Gene: 32
C;Superfamily: herpesvirus infected cell protein ICP18.5
C;Keywords: capsid assembly

Query Match 1.9%; Score 8; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 7.8;

```
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     629 LKEELARF 636


RESULT   5
T42575
infected cell protein 32 - equine herpesvirus 4 (strain NS80567)
C;Species: equine herpesvirus 4
A;Variety: strain NS80567
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C;Accession: T42575
R;Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.
J. Gen. Virol. 79, 1197-1203, 1998
A;Title: The DNA sequence of equine herpesvirus-4.
A;Reference number: Z22173; MUID:98264497
A;Accession: T42575
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-773 <TEL>
A;Cross-references: EMBL:AF030027; NID:g2605950; PIDN:AAC59548.1; PID:g2605976
A;Experimental source: strain NS80567
C;Genetics:
A;Gene: 32
C;Superfamily: herpesvirus infected cell protein ICP18.5
C;Keywords: capsid assembly


  Query Match             1.9%;  Score 8;  DB 2;  Length 773;
  Best Local Similarity   100.0%;  Pred. No. 7.9;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     636 LKEELARF 643


RESULT   6
WMBEH8
infected cell protein ICP18.5 - human herpesvirus 1 (strain F)
C;Species: human herpesvirus 1
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C;Accession: A26101
R;Pellett, P.E.; Jenkins, F.J.; Ackermann, M.; Sarmiento, M.; Roizman, B.
J. Virol. 60, 1134-1140, 1986
A;Title: Transcription initiation sites and nucleotide sequence of a herpes simplex viru
A;Reference number: A26101; MUID:87061190
A;Accession: A26101
A;Molecule type: DNA
A;Residues: 1-780 <PEL>
A;Cross-references: GB:M14164; GB:M12398; NID:g330084; PIDN:AAA45775.1; PID:g330085
C;Comment: This protein may affect translocation of herpes simplex virus glycoproteins t
C;Genetics:
A;Gene: ICP18.5
C;Superfamily: herpesvirus infected cell protein ICP18.5
C;Keywords: capsid assembly


  Query Match             1.9%;  Score 8;  DB 1;  Length 780;
  Best Local Similarity   100.0%;  Pred. No. 8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     635 LKEELARF 642


RESULT   7
WMBEK8
infected cell protein ICP18.5 - human herpesvirus 1 (strain Angelotti)
C;Species: human herpesvirus 1
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jul-1999
C;Accession: A26987
R;Knopf, C.W.
Nucleic Acids Res. 15, 8109-8110, 1987
A;Title: The nucleotide sequence of the herpes simplex virus type 1 late gene ICP18.5
A;Reference number: A26987; MUID:88040433
A;Accession: A26987
A;Molecule type: DNA
A;Residues: 1-785 <KNO>
A;Cross-references: GB:Y00453; NID:g60416; PIDN:CAA68508.1; PID:g60417
C;Comment: This protein is involved in the translocation of viral glycoproteins to me
C;Superfamily: herpesvirus infected cell protein ICP18.5
C;Keywords: capsid assembly; late protein


  Query Match             1.9%;  Score 8;  DB 1;  Length 785;
  Best Local Similarity   100.0%;  Pred. No. 8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     640 LKEELARF 647


RESULT   8
WMBEW8
infected cell protein ICP18.5 - human herpesvirus 1 (strain 17)
C;Species: human herpesvirus 1
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jun-2000
C;Accession: A30085
R;McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; P
J. Gen. Virol. 69, 1531-1574, 1988
A;Title: The complete DNA sequence of the long unique region in the genome of herpes
A;Reference number: A30083; MUID:88274327
A;Accession: A30085
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-785 <MCG>
A;Cross-references: GB:X14112; NID:g1944536; PIDN:CAA32321.1; PID:g59528; GB:D00317
C;Superfamily: herpesvirus infected cell protein ICP18.5
C;Keywords: capsid assembly; late protein


  Query Match             1.9%;  Score 8;  DB 1;  Length 785;
  Best Local Similarity   100.0%;  Pred. No. 8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     640 LKEELARF 647


RESULT   9
T05501
hypothetical protein T19K4.210 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C;Accession: T05501
R;Bevan, M.; Wedler, H.; Wambutt, R.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.X.; Schue
submitted to the Protein Sequence Database, April 1998
A;Reference number: Z15418
A;Accession: T05501
A;Molecule type: DNA
A;Residues: 1-3738 <BEV>
A;Cross-references: EMBL:AL022373
A;Experimental source: cultivar Columbia; BAC clone T19K4
C;Genetics:
A;Map position: 4
A;Introns: 20/1; 114/2; 243/3; 289/1; 333/2; 391/3; 412/3; 432/3; 438/3; 460/3; 482/1
68/3; 2152/3; 2168/3; 2298/2; 2446/3; 2879/3; 2966/3; 3029/3; 3049/3; 3183/2; 3352/3;
```

```
A;Note: T19K4.210


  Query Match             1.9%;  Score 8;  DB 2;  Length 3738;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     307 KVLTSLNR 314
           ||||||||
Db    1312 KVLTSLNR 1319


RESULT  10
T00831
hypothetical protein At2g17930 [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein T13L16.5
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 02-Feb-2001
C;Accession: T00831; B84558
R;de la Bastide, M.; Hameed, A.; Gnoj, L.; Jensen, K.; Shohdy, N.; Gottesman, T.; Haberm
 McCombie, W.R.
submitted to the EMBL Data Library, January 1999
A;Description: A. thaliana BAC T13L16 from chromosome IV, top arm.
A;Reference number: Z14205
A;Accession: T00831
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-3795 <DEL>
A;Cross-references: EMBL:AC003952; NID:g2708736; PID:g2708741
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487
A;Accession: B84558
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-3795 <STO>
A;Cross-references: GB:AE002093; NID:g6598526; PIDN:AAD20114.2; GSPDB:GN00139
C;Genetics:
A;Gene: At2g17930; T13L16.5
A;Map position: 2
A;Introns: 20/1; 258/3; 287/2; 323/3; 355/3; 376/3; 396/3; 419/3; 441/1; 458/3; 467/3; 8
938/3; 3025/3; 3122/3; 3265/3; 3409/3; 3502/3; 3686/3; 3756/3


  Query Match             1.9%;  Score 8;  DB 2;  Length 3795;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     307 KVLTSLNR 314
           ||||||||
Db    1258 KVLTSLNR 1265


RESULT  11
T14687
hypothetical protein - Yersinia pestis plasmid pMT1
C;Species: Yersinia pestis
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14687
R;Hu, P.; Elliott, J.; McCready, P.; Skowronski, E.; Garnes, J.; Kobayashi, A.; Carrano,
submitted to the EMBL Data Library, March 1998
A;Description: Structural organization of virulence determinants in three Yersinia pesti
A;Reference number: Z18168
A;Accession: T14687
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-74 <HUP>
A;Cross-references: EMBL:AF053947; NID:g2996286; PID:g2996322; PIDN:AAC13202.1
C;Genetics:
A;Genome: plasmid pMT1


  Query Match             1.6%;  Score 7;  DB 2;  Length 74;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     394 LTAIYIR 400
           |||||||
Db      27 LTAIYIR 33


RESULT  12
C84094
protein translocation protein BH3555 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
C;Accession: C84094
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A;Reference number: A83650; MUID:20263314
A;Accession: C84094
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-76 <STO>
A;Cross-references: GB:AP001519; GB:BA000004; NID:g10176109; PIDN:BAB07274.1; GSPDB:G
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3555


  Query Match             1.6%;  Score 7;  DB 2;  Length 76;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     233 FVLTIAV 239
           |||||||
Db      66 FVLTIAV 72


RESULT  13
M3HUGA
Ig heavy chain V-III region (Ga) - human (tentative sequence)
C;Species: Homo sapiens (man)
C;Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 31-Mar-2000
C;Accession: A02052
R;Florent, G.; Lehman, D.; Putnam, F.W.
Biochemistry 13, 2482-2498, 1974
A;Title: The switch point in mu heavy chains of human IgM immunoglobulins.
A;Reference number: A02052; MUID:74175307
A;Accession: A02052
A;Molecule type: protein
A;Residues: 1-122 <FLO>
C;Comment: This chain was isolated from a Waldenstrom's macroglobulin.
C;Genetics:
A;Gene: GDB:IGHV@
A;Cross-references: GDB:128528; OMIM:147070
A;Map position: 14q32.33-14q32.33
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin; pyroglutamic acid
F;15-98/Domain: immunoglobulin homology <IMM>
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;22-96/Disulfide bonds: #status predicted


  Query Match             1.6%;  Score 7;  DB 1;  Length 122;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     297 WLSVISY 303
           |||||||
Db      47 WLSVISY 53
```

RESULT 14
E82653
hypothetical protein XF1666 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: E82653
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: E82653
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-137 <SIM>
A;Cross-references: GB:AE003992; GB:AE003849; NID:g9106715; PIDN:AAF84475.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laign
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1666

Query Match 1.6%; Score 7; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 111 LARFKPY 117
|||||||
Db 92 LARFKPY 98

RESULT 15
S25660
4-hydroxybenzoate synthetase - Escherichia coli
N;Alternate names: chorismate lyase; chorismate pyruvate-lyase
C;Species: Escherichia coli
C;Date: 22-Nov-1993 #sequence_revision 21-Jul-1995 #text_change 04-Mar-2000
C;Accession: S25660; A42956; S24360; I70800; F65211
R;Nishimura, K.; Nakahigashi, K.; Inokuchi, H.
J. Bacteriol. 174, 5762, 1992
A;Title: Location of the ubiA gene on the physical map of Escherichia coli.
A;Reference number: S25660; MUID:92380960
A;Accession: S25660
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-202 <NIS>
A;Cross-references: EMBL:X57434; NID:g43233; PIDN:CAA40681.1; PID:g43234
A;Experimental source: strain K12 W3110
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1991
R;Nichols, B.P.; Green, J.M.
J. Bacteriol. 174, 5309-5316, 1992
A;Title: Cloning and sequencing of Escherichia coli ubiC and purification of chorismate
A;Reference number: A42956; MUID:92355505
A;Accession: A42956
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 38-202 <NIC>
A;Cross-references: GB:M93413; NID:g148106; PIDN:AAA24716.1; PID:g148107
A;Note: sequence extracted from NCBI backbone (NCBIP:110479)
R;Siebert, M.; Bechthold, A.; Melzer, M.; May, U.; Berger, U.; Schroeder, G.; Schroed
FEBS Lett. 307, 347-350, 1992
A;Title: Ubiquinone biosynthesis. Cloning of the genes coding for chorismate pyruvate
A;Reference number: S24360; MUID:92354744
A;Accession: S24360
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 38-202 <SIE>
A;Cross-references: EMBL:X66619; NID:g43230; PIDN:CAA47181.1; PID:g43231
R;Wu, G.; Williams, H.D.; Gibson, F.; Poole, R.K.
J. Gen. Microbiol. 139, 1795-1805, 1993
A;Title: Mutants of Escherichia coli affected in respiration: the cloning and nucleot
A;Reference number: I55717; MUID:94014977
A;Accession: I70800
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 38-202 <RES>
A;Cross-references: GB:M96268; NID:g347886; PIDN:AAA17027.1; PID:g347888
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617
A;Accession: F65211
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-202 <BLAT>
A;Cross-references: GB:AE000477; GB:U00096; NID:g2367338; PIDN:AAC77009.1; PID:g17904
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ubiC
C;Superfamily: Escherichia coli 4-hydroxybenzoate synthetase

Query Match 1.6%; Score 7; DB 2; Length 202;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 321 FLRYNAF 327
|||||||
Db 23 FLRYNAF 29

Search completed: October 30, 2001, 15:30:51
Job time: 267 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 15:30:14 ; Search time 10.78 Seconds
(without alignments)
1353.695 Million cell updates/sec

Title: US-09-619-049-855
Perfect score: 426
Sequence: 1 MRLPWVIFCTVLLLIFTNNS.....NKFTNEEEGRINCINFDSQT 426

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	8	1.9	664	1	PRTP_HSVB2	P17587 bovine herp
2	8	1.9	724	1	PRTP_PRVIF	P11871 pseudorabie
3	8	1.9	766	1	PRTP_HSVEB	P28973 equine herp
4	8	1.9	780	1	PRTP_HSV1F	P06490 herpes simp
5	8	1.9	785	1	PRTP_HSV11	P10212 herpes simp
6	8	1.9	785	1	PRTP_HSV1A	P12835 herpes simp
7	7	1.6	122	1	HV3H_HUMAN	P01769 homo sapien
8	7	1.6	262	1	DRN1_PIG	P11936 sus scrofa
9	7	1.6	307	1	ALF_HELPY	P56109 helicobacte
10	7	1.6	344	1	DHB1_MOUSE	P51656 mus musculu
11	7	1.6	344	1	DHB1_RAT	P51657 rattus norv
12	7	1.6	353	1	PERM_ECOLI	P77406 escherichia
13	7	1.6	380	1	SUC2_ARCFU	O28097 archaeoglob
14	7	1.6	399	1	K1CS_BOVIN	P08728 bos taurus
15	7	1.6	400	1	K1CS_HUMAN	P08727 homo sapien
16	7	1.6	401	1	K1CS_POTTR	P51856 potorous tr
17	7	1.6	430	1	CGB1_MOUSE	P24860 mus musculu
18	7	1.6	469	1	K1CX_HUMAN	P08779 homo sapien
19	7	1.6	473	1	K1CP_HUMAN	P30654 homo sapien
20	7	1.6	533	1	THAS_RAT	P49430 rattus norv
21	7	1.6	651	1	YHJK_ECOLI	P37649 escherichia
22	7	1.6	731	1	MASZ_MYCLE	O32913 mycobacteri
23	7	1.6	741	1	MASZ_MYCTU	Q50596 mycobacteri
24	7	1.6	937	1	DBS_RAT	Q63406 rattus norv
25	7	1.6	1108	1	DBS_HUMAN	O15068 homo sapien
26	7	1.6	1149	1	DBS_MOUSE	Q64096 mus musculu
27	7	1.6	1206	1	RPOC_BACHD	Q9z9m1 bacillus ha
28	7	1.6	1273	1	WEB1_YEAST	P38968 saccharomyc
29	6	1.4	19	1	LPGE_ECOLI	P33236 escherichia
30	6	1.4	38	1	PSBX_ODOSI	P49509 odontella s
31	6	1.4	42	1	YHJX_SALTY	P43665 salmonella
32	6	1.4	56	1	SSPF_BACME	P52971 bacillus me
33	6	1.4	67	1	TATA_NEIMA	P57049 neisseria m
34	6	1.4	76	1	SECG_BACSU	O32233 bacillus su
35	6	1.4	87	1	XHLB_BACSU	Q99163 bacillus su
36	6	1.4	91	1	MERP_SHEPU	Q54463 shewanella
37	6	1.4	92	1	DBH_BACSU	P08821 bacillus su
38	6	1.4	97	1	VE7_HPV55	Q80935 human papil
39	6	1.4	102	1	HBAZ_MESAU	P29627 mesocricetu
40	6	1.4	108	1	Y695_METJA	Q58106 methanococc
41	6	1.4	111	1	VP34_BPAPS	Q9t1r4 bacteriopha
42	6	1.4	118	1	RS13_BUCAI	P57569 buchnera ap
43	6	1.4	121	1	RS13_HAEIN	P44380 haemophilus
44	6	1.4	124	1	YEA3_SCHPO	O14068 schizosacch
45	6	1.4	139	1	PRO2_HUMAN	P35080 homo sapien
46	6	1.4	142	1	TOLR_ECOLI	P05829 escherichia
47	6	1.4	146	1	HBB0_MOUSE	P04443 mus musculu
48	6	1.4	146	1	HBBY_MESAU	P29626 mesocricetu
49	6	1.4	146	1	HBBZ_MOUSE	P04444 mus musculu
50	6	1.4	146	1	HBG1_PONPY	P18995 pongo pygma
51	6	1.4	146	1	HBG2_PONPY	P18996 pongo pygma
52	6	1.4	146	1	HBG_ALOBE	P56284 alouatta be
53	6	1.4	146	1	HBG_ALOSE	P56285 alouatta se
54	6	1.4	146	1	HBG_ATEGE	P06891 ateles geof
55	6	1.4	146	1	HBG_CEBAP	Q28225 cebus apell
56	6	1.4	146	1	HBG_CHEME	P08224 cheirogaleu
57	6	1.4	146	1	HBG_EULFU	P08225 eulemur ful
58	6	1.4	146	1	HBG_GORGO	P06641 gorilla gor
59	6	1.4	146	1	HBG_HUMAN	P02096 homo sapien
60	6	1.4	146	1	HBG_HYLLA	P16175 hylobates l
61	6	1.4	146	1	HBG_MACMU	P02098 macaca mula
62	6	1.4	146	1	HBG_MACNE	P02097 macaca neme
63	6	1.4	149	1	FUR_HAEDU	P71333 haemophilus
64	6	1.4	151	1	CALM_DICDI	P02599 dictyosteli
65	6	1.4	151	1	CALM_PNECA	P41041 pneumocysti
66	6	1.4	158	1	SH3B_DROME	Q9nfp5 drosophila
67	6	1.4	159	1	Y207_MYCPN	P75349 mycoplasma
68	6	1.4	163	1	RS19_AERPE	Q9yf74 aeropyrum p
69	6	1.4	167	1	YBL7_SCHPO	Q9urv4 schizosacch
70	6	1.4	169	1	FLAV_ANASP	P11241 anabaena sp
71	6	1.4	170	1	CBP4_YEAST	P37267 saccharomyc
72	6	1.4	178	1	YC96_HAEIN	Q57519 haemophilus
73	6	1.4	181	1	APT2_YEAST	P36973 saccharomyc
74	6	1.4	188	1	GSPH_ERWCA	P31587 erwinia car
75	6	1.4	193	1	TNF7_HUMAN	P32970 homo sapien
76	6	1.4	198	1	NQRE_VIBAL	Q56589 vibrio algi
77	6	1.4	200	1	GCAP_HUMAN	P43080 homo sapien
78	6	1.4	201	1	GCAP_MOUSE	P43081 mus musculu
79	6	1.4	204	1	GCAP_BOVIN	P46065 bos taurus
80	6	1.4	205	1	TF2D_DICDI	P26355 dictyosteli
81	6	1.4	207	1	Y388_METJA	Q57833 methanococc
82	6	1.4	213	1	GS27_CAEEL	P41941 caenorhabdi
83	6	1.4	215	1	YCEH_ECOLI	P29217 escherichia
84	6	1.4	222	1	R11C_TOBAC	Q40520 nicotiana t
85	6	1.4	222	1	R11D_TOBAC	Q40522 nicotiana t
86	6	1.4	222	1	Y4YL_RHISN	P55720 rhizobium s
87	6	1.4	229	1	YGE5_YEAST	P53180 saccharomyc
88	6	1.4	234	1	GU33_RAT	P35896 rattus norv
89	6	1.4	234	1	YIV8_YEAST	P40582 saccharomyc
90	6	1.4	236	1	IPT_AGRVI	Q04590 agrobacteri
91	6	1.4	237	1	LECA_DOLLA	P38662 dolichos la
92	6	1.4	240	1	SDIA_ECOLI	P07026 escherichia
93	6	1.4	241	1	GTO1_PIG	Q9n1f5 sus scrofa
94	6	1.4	242	1	MYF6_CHICK	Q01795 gallus gall
95	6	1.4	242	1	MYF6_HUMAN	P23409 homo sapien
96	6	1.4	242	1	MYF6_MOUSE	P15375 mus musculu
97	6	1.4	243	1	OPAD_NEIGO	Q04883 neisseria g
98	6	1.4	244	1	OPAI_NEIGO	Q04877 neisseria g
99	6	1.4	245	1	Y474_TREPA	O83487 treponema p
100	6	1.4	248	1	Y964_PSEAE	Q51423 pseudomonas
101	6	1.4	255	1	K1CS_RAT	Q63279 rattus norv
102	6	1.4	256	1	RS3A_CAEEL	P48154 caenorhabdi
103	6	1.4	259	1	OVUH_LYMST	P06308 lymnaea sta
104	6	1.4	264	1	PANB_ECOLI	P31057 escherichia
105	6	1.4	266	1	RBM7_HUMAN	Q9y580 homo sapien
106	6	1.4	267	1	RS3A_DROME	P55830 drosophila

107	6	1.4	268	1	RS3A_ANOGA	P52813	anopheles g	180	6	1.4	394	1	SETC_ECOLI	P31436	escherichia
108	6	1.4	269	1	COX3_YEAST	P00420	saccharomyc	181	6	1.4	396	1	PGLR_TOBAC	Q05967	nicotiana t
109	6	1.4	270	1	OMPC_NEIGO	P09888	neisseria g	182	6	1.4	396	1	YJIB_BACSU	O34374	bacillus su
110	6	1.4	271	1	ISPE_THEMA	Q9x1a3	thermotoga	183	6	1.4	397	1	A2AB_TALEU	O19091	talpa europ
111	6	1.4	274	1	YDDG_ECOLI	P46136	escherichia	184	6	1.4	397	1	ASSY_STRCL	P50986	streptomyce
112	6	1.4	279	1	GR78_PLAFA	P12794	plasmodium	185	6	1.4	402	1	YHJX_ECOLI	P37662	escherichia
113	6	1.4	286	1	NIFC_CLOPA	P18795	clostridium	186	6	1.4	403	1	K1CS_MOUSE	P19001	mus musculu
114	6	1.4	288	1	YDIB_ECOLI	P28244	escherichia	187	6	1.4	406	1	TYRQ_HAEIN	P44747	haemophilus
115	6	1.4	291	1	CDK2_ENTHI	Q04770	entamoeba h	188	6	1.4	410	1	CREA_BACSP	P38487	bacillus sp
116	6	1.4	293	1	RFFH_ECOLI	P27831	escherichia	189	6	1.4	410	1	YB07_FOWPM	P14360	fowlpox vir
117	6	1.4	295	1	NMCR_ENTCL	P52676	enterobacte	190	6	1.4	414	1	IDHC_HUMAN	O75874	homo sapien
118	6	1.4	305	1	E2BA_CAEEL	P34604	caenorhabdi	191	6	1.4	415	1	TRSA_STRAM	Q07197	streptomyce
119	6	1.4	307	1	ALF_HELPJ	Q9zmq6	helicobacte	192	6	1.4	416	1	SYH_METJA	Q58406	methanococc
120	6	1.4	310	1	FTSX_HAEIN	P44872	haemophilus	193	6	1.4	417	1	FRZC_MYXXA	P43500	myxococcus
121	6	1.4	310	1	HEMZ_BACHD	Q9kdk9	bacillus ha	194	6	1.4	422	1	NCAP_VSVJO	P04881	vesicular s
122	6	1.4	311	1	ATH6_ARATH	P46668	arabidopsis	195	6	1.4	428	1	DCTA_ECOLI	P37312	escherichia
123	6	1.4	311	1	TRXB_HELPJ	Q9zl18	helicobacte	196	6	1.4	428	1	DCTA_SALTY	P50334	salmonella
124	6	1.4	311	1	TRXB_HELPY	P56431	helicobacte	197	6	1.4	428	1	EF1A_METVA	P07810	methanococc
125	6	1.4	312	1	GU27_RAT	P34987	rattus norv	198	6	1.4	429	1	CGB1_CRILO	Q08301	cricetulus
126	6	1.4	314	1	AA3R_CANFA	Q28309	canis famil	199	6	1.4	431	1	K1CQ_HUMAN	Q04695	homo sapien
127	6	1.4	314	1	KPRS_ECOLI	P08330	escherichia	200	6	1.4	432	1	A2AR_LABOS	Q91081	labrus ossi
128	6	1.4	314	1	KPRS_SALTY	P15849	salmonella	201	6	1.4	435	1	VGLM_VZVD	P09298	varicella-z
129	6	1.4	315	1	KPRS_BUCAI	P57266	buchnera ap	202	6	1.4	435	1	YYBO_BACSU	P37489	bacillus su
130	6	1.4	315	1	KPRS_HAEIN	P44328	haemophilus	203	6	1.4	436	1	THCB_RHOER	P43492	rhodococcus
131	6	1.4	318	1	AA3R_HUMAN	P33765	homo sapien	204	6	1.4	448	1	A2AB_CAVPO	Q60475	cavia porce
132	6	1.4	320	1	YFHA_BORPE	P33445	bordetella	205	6	1.4	450	1	A2AA_CAVPO	Q60474	cavia porce
133	6	1.4	322	1	NU1M_PELSU	O79670	pelomedusa	206	6	1.4	450	1	A2AA_HUMAN	P08913	homo sapien
134	6	1.4	322	1	RLA0_LEICH	P39096	leishmania	207	6	1.4	450	1	A2AA_MOUSE	Q01338	mus musculu
135	6	1.4	323	1	RLA0_LEIIN	P39097	leishmania	208	6	1.4	450	1	A2AA_PIG	P18871	sus scrofa
136	6	1.4	324	1	TCH3_ARATH	P25071	arabidopsis	209	6	1.4	450	1	A2AA_RAT	P22909	rattus norv
137	6	1.4	328	1	HOLB_PSEAE	P52024	pseudomonas	210	6	1.4	450	1	A2AB_HUMAN	P18089	homo sapien
138	6	1.4	331	1	UL38_HCMVA	P16779	human cytom	211	6	1.4	452	1	A2AA_BOVIN	Q28838	bos taurus
139	6	1.4	333	1	CATJ_MOUSE	Q9r014	mus musculu	212	6	1.4	453	1	2ABD_RAT	P56932	r serine/th
140	6	1.4	333	1	DPPB_HAEIN	P45096	haemophilus	213	6	1.4	453	1	A2AB_RAT	P19328	rattus norv
141	6	1.4	335	1	FABH_CHLPN	Q9z8p0	chlamydia p	214	6	1.4	455	1	A2AB_MOUSE	P30545	mus musculu
142	6	1.4	335	1	UL16_HSV6U	P24442	human herpe	215	6	1.4	456	1	ACHA_CHICK	P09479	gallus gall
143	6	1.4	338	1	G3P_METBR	P19314	methanobact	216	6	1.4	458	1	NU4M_PETMA	Q35542	petromyzon
144	6	1.4	340	1	VB17_VACCC	P21075	vaccinia vi	217	6	1.4	459	1	NU4M_POLOR	Q95917	polypterus
145	6	1.4	340	1	VB17_VARV	P33878	variola vir	218	6	1.4	460	1	NU4M_CARAU	O78687	carassius a
146	6	1.4	349	1	END3_YEAST	P39013	saccharomyc	219	6	1.4	460	1	NU4M_CROLA	P34194	crossostoma
147	6	1.4	350	1	RFBB_RHISN	P55462	rhizobium s	220	6	1.4	460	1	NU4M_ONCMY	P11631	oncorhynchu
148	6	1.4	354	1	US28_HCMVA	P09704	human cytom	221	6	1.4	460	1	NU4M_SALSA	Q9zzm4	salmo salar
149	6	1.4	356	1	GPDA_TREPA	O83973	treponema p	222	6	1.4	461	1	NCB1_HUMAN	Q02818	homo sapien
150	6	1.4	356	1	NTRB_RHOCA	P09431	rhodobacter	223	6	1.4	461	1	NU4M_LATCH	O03173	latimeria c
151	6	1.4	359	1	K1CN_MOUSE	Q61781	mus musculu	224	6	1.4	467	1	K1C1_CARAU	Q90303	carassius a
152	6	1.4	362	1	AROB_BACSU	P31102	bacillus su	225	6	1.4	471	1	IAAG_MAIZE	Q41819	zea mays (m
153	6	1.4	364	1	GHSR_RAT	O08725	rattus norv	226	6	1.4	471	1	K1CN_HUMAN	P02533	homo sapien
154	6	1.4	365	1	SUCC_METTH	O27115	methanobact	227	6	1.4	475	1	CO2B_HUMAN	Q9uq03	homo sapien
155	6	1.4	366	1	GHSR_HUMAN	Q92847	homo sapien	228	6	1.4	475	1	DXR_MENPI	Q9xes0	mentha pipe
156	6	1.4	369	1	YAD6_YEAST	P39729	saccharomyc	229	6	1.4	477	1	DXR_ARATH	Q9xfs9	arabidopsis
157	6	1.4	372	1	MK11_HUMAN	Q15759	homo sapien	230	6	1.4	480	1	ASSY_STRCO	P24532	streptomyce
158	6	1.4	376	1	EXD_DROME	P40427	drosophila	231	6	1.4	481	1	ASSY_STRLA	Q06734	streptomyce
159	6	1.4	376	1	Y550_BUCAI	P57615	buchnera ap	232	6	1.4	483	1	PEPD_HAEIN	P44817	haemophilus
160	6	1.4	378	1	AI3M_YEAST	P03877	saccharomyc	233	6	1.4	486	1	ZPR1_YEAST	P53303	saccharomyc
161	6	1.4	379	1	YLEU_DEBOC	P48012	debaryomyce	234	6	1.4	487	1	AMPL_BOVIN	P00727	bos taurus
162	6	1.4	382	1	A2AB_DIDMA	O77715	didelphis m	235	6	1.4	487	1	CPDE_BOVIN	Q01361	bos taurus
163	6	1.4	383	1	NRL1_RHORH	Q02068	rhodococcus	236	6	1.4	489	1	UDPE_NPVOP	Q65363	orgyia pseu
164	6	1.4	384	1	A2AB_ECHTE	O77723	echinops te	237	6	1.4	490	1	CPC7_RAT	P05179	rattus norv
165	6	1.4	384	1	A2AB_ELEMA	O19014	elephas max	238	6	1.4	495	1	DNAA_MYCSM	P49992	mycobacteri
166	6	1.4	385	1	RECF_MYCLE	P46391	mycobacteri	239	6	1.4	501	1	TRA2_MOUSE	P39429	mus musculu
167	6	1.4	386	1	A2AB_AMBHO	O18935	amblysomus	240	6	1.4	503	1	CPV1_HORSE	O46512	equus cabal
168	6	1.4	386	1	YD46_MYCTU	Q11016	mycobacteri	241	6	1.4	503	1	IRF7_HUMAN	Q92985	homo sapien
169	6	1.4	387	1	A2AB_MACPR	O19025	macroscelid	242	6	1.4	504	1	GAG_HV1MA	P04594	human immun
170	6	1.4	387	1	PKNS_MYCGE	P47355	mycoplasma	243	6	1.4	504	1	GAL1_CANPA	O42821	candida par
171	6	1.4	387	1	VPD_BPP2	P10312	bacteriopha	244	6	1.4	507	1	SXA2_SCHPO	P32825	schizosacch
172	6	1.4	388	1	A2AB_ORYAF	O19032	orycteropus	245	6	1.4	509	1	CPT7_PIG	P19100	sus scrofa
173	6	1.4	389	1	A2AB_PROHA	O19054	procavia ca	246	6	1.4	509	1	KPYK_SCHPO	Q10208	schizosacch
174	6	1.4	389	1	VPD_BP186	P21679	bacteriopha	247	6	1.4	519	1	RHO_TREPA	O83281	treponema p
175	6	1.4	390	1	A2AB_DUGDU	O77713	dugong dugo	248	6	1.4	522	1	INX8_CAEEL	Q23593	caenorhabdi
176	6	1.4	391	1	A2AB_ERIEU	O19012	erinaceus e	249	6	1.4	522	1	NIFK_AZOVI	P07329	azotobacter
177	6	1.4	392	1	A2AB_BOVIN	O77700	bos taurus	250	6	1.4	527	1	SECY_PEA	Q9xqu4	pisum sativ
178	6	1.4	393	1	FTSW_SYNY3	P74180	synechocyst	251	6	1.4	532	1	KPYK_AGABI	O94122	agaricus bi
179	6	1.4	394	1	A2AB_RABIT	O77830	oryctolagus	252	6	1.4	534	1	GCR2_YEAST	Q01722	saccharomyc

253	6	1.4	534	1	K2C4_HUMAN	P19013	homo sapien
254	6	1.4	543	1	PKAA_STRCO	P54739	streptomyce
255	6	1.4	545	1	SECY_SPIOL	P93690	spinacia ol
256	6	1.4	546	1	CG11_YEAST	P20437	saccharomyc
257	6	1.4	554	1	MASZ_YEAST	P21826	saccharomyc
258	6	1.4	555	1	O16G_BACCO	Q45101	bacillus co
259	6	1.4	560	1	HEMA_IATKR	P03458	influenza a
260	6	1.4	560	1	YDEN_ECOLI	P77318	escherichia
261	6	1.4	566	1	KAD5_SCHPO	Q09831	schizosacch
262	6	1.4	569	1	RS1_CHLTR	O84100	chlamydia t
263	6	1.4	570	1	RS1_CHLMU	P38016	chlamydia m
264	6	1.4	573	1	CH60_DROME	O02649	drosophila
265	6	1.4	574	1	5NTD_HUMAN	P21589	homo sapien
266	6	1.4	580	1	RS1_CHLPN	Q9z8m3	chlamydia p
267	6	1.4	587	1	FIBP_ADE12	P36711	human adeno
268	6	1.4	590	1	PRIM_CHLPN	Q9z6w4	chlamydia p
269	6	1.4	594	1	MG11_MOUSE	Q60710	mus musculu
270	6	1.4	602	1	NU5M_MACRO	P92669	macropus ro
271	6	1.4	603	1	US26_HCMVA	P09699	human cytom
272	6	1.4	609	1	GPI1_YEAST	P53306	saccharomyc
273	6	1.4	612	1	NU5M_CROLA	P34195	crossostoma
274	6	1.4	616	1	TYPA_HAEIN	P44910	haemophilus
275	6	1.4	618	1	PDA4_CAEEL	P34329	caenorhabdi
276	6	1.4	623	1	NRC2_NEUCR	O42626	neurospora
277	6	1.4	627	1	CACP_CANTR	Q00614	candida tro
278	6	1.4	628	1	YKP5_YEAST	P36056	saccharomyc
279	6	1.4	630	1	HCYA_EURCA	P14750	eurypelma c
280	6	1.4	630	1	YIB4_YEAST	P40549	saccharomyc
281	6	1.4	632	1	NTG3_HUMAN	P48066	homo sapien
282	6	1.4	635	1	NRFE_HAEIN	P44944	haemophilus
283	6	1.4	638	1	YD14_SCHPO	Q10237	schizosacch
284	6	1.4	641	1	DXS_RHOCA	P26242	rhodobacter
285	6	1.4	643	1	RHOP_MOUSE	Q61085	mus musculu
286	6	1.4	647	1	MTF1_FLAOK	P14871	flavobacter
287	6	1.4	649	1	YAY3_SCHPO	Q10211	schizosacch
288	6	1.4	654	1	CBP1_YEAST	P07252	saccharomyc
289	6	1.4	655	1	GR78_PLAFO	Q05866	plasmodium
290	6	1.4	657	1	GLGX_ECOLI	P15067	escherichia
291	6	1.4	658	1	CNT1_RABIT	Q9mzt2	oryctolagus
292	6	1.4	667	1	OS9_HUMAN	Q13438	homo sapien
293	6	1.4	675	1	NU5M_ACACA	Q37372	acanthamoeb
294	6	1.4	676	1	EXL1_HUMAN	Q92935	homo sapien
295	6	1.4	676	1	QR1_COTJA	P23499	coturnix co
296	6	1.4	677	1	NCPR_MOUSE	P37040	mus musculu
297	6	1.4	677	1	NCPR_RAT	P00388	rattus norv
298	6	1.4	707	1	KPC2_CAEEL	P34885	caenorhabdi
299	6	1.4	708	1	PTA_BUCAI	P57273	buchnera ap
300	6	1.4	710	1	PAL2_ORYSA	P53443	oryza sativ
301	6	1.4	720	1	HS78_SCHPO	O59838	schizosacch
302	6	1.4	726	1	PLO1_BOVIN	O77588	bos taurus
303	6	1.4	727	1	PLO1_HUMAN	Q02809	homo sapien
304	6	1.4	728	1	PLO1_RAT	Q63321	rattus norv
305	6	1.4	730	1	PLO1_CHICK	P24802	gallus gall
306	6	1.4	732	1	COPA_HELFE	O32619	helicobacte
307	6	1.4	738	1	MASZ_CORGL	P42450	corynebacte
308	6	1.4	743	1	STB5_YEAST	P38699	saccharomyc
309	6	1.4	747	1	DIL2_HUMAN	Q9ulw0	homo sapien
310	6	1.4	748	1	Y875_METJA	Q58285	methanococc
311	6	1.4	756	1	FILS_BOVIN	Q06002	bos taurus
312	6	1.4	757	1	MUS2_THEMA	Q9x105	thermotoga
313	6	1.4	760	1	RESA_PLAFN	P13831	plasmodium
314	6	1.4	761	1	P52K_HUMAN	O43422	homo sapien
315	6	1.4	770	1	PRTP_VZVD	P09284	varicella-z
316	6	1.4	783	1	LON_VIBPA	P74956	vibrio para
317	6	1.4	784	1	LON_ECOLI	P08177	escherichia
318	6	1.4	784	1	LON_ERWAM	P46067	erwinia amy
319	6	1.4	789	1	LP1B_DROME	P11996	drosophila
320	6	1.4	796	1	YEZ9_YEAST	P39956	saccharomyc
321	6	1.4	806	1	HYF2_ALCEU	O07451	alcaligenes
322	6	1.4	810	1	PLSB_HAEIN	P44857	haemophilus
323	6	1.4	821	1	YNZ0_CAEEL	P45970	caenorhabdi
324	6	1.4	825	1	QUTA_EMENI	P10563	emericella
325	6	1.4	826	1	HIFA_HUMAN	Q16665	homo sapien
326	6	1.4	829	1	GYRA_CLOAB	P94605	clostridium
327	6	1.4	831	1	YMX3_CAEEL	P34511	caenorhabdi
328	6	1.4	834	1	GYRA_CHLPN	Q9z8r4	chlamydia p
329	6	1.4	837	1	ROD1_YEAST	Q02805	saccharomyc
330	6	1.4	843	1	MNB_DROME	P49657	drosophila
331	6	1.4	846	1	YDSB_SCHPO	O14186	schizosacch
332	6	1.4	848	1	TORA_ECOLI	P33225	escherichia
333	6	1.4	855	1	ENV_FIVU8	Q04995	feline immu
334	6	1.4	857	1	PAX6_DROME	O18381	drosophila
335	6	1.4	859	1	OSH1_YEAST	P35845	saccharomyc
336	6	1.4	905	1	HEX_ADECR	O39619	canine aden
337	6	1.4	906	1	GLR1_HUMAN	P42261	homo sapien
338	6	1.4	907	1	GLR1_MOUSE	P23818	mus musculu
339	6	1.4	907	1	GLR1_RAT	P19490	rattus norv
340	6	1.4	925	1	DBL_HUMAN	P10911	homo sapien
341	6	1.4	926	1	ATCL_SYNP7	P37278	synechococc
342	6	1.4	956	1	YEF3_YEAST	P32618	saccharomyc
343	6	1.4	968	1	Y682_METJA	Q58095	methanococc
344	6	1.4	976	1	XPR6_YARLI	P42781	yarrowia li
345	6	1.4	1012	1	POLG_IBDVO	P27276	avian infec
346	6	1.4	1013	1	A60D_DROME	P91927	drosophila
347	6	1.4	1021	1	MANA_RHOMR	P49425	rhodothermu
348	6	1.4	1029	1	YFC5_YEAST	P43571	saccharomyc
349	6	1.4	1042	1	SYI_BORBU	O51773	borrelia bu
350	6	1.4	1047	1	MSH3_YEAST	P25336	saccharomyc
351	6	1.4	1063	1	SPT5_YEAST	P27692	saccharomyc
352	6	1.4	1073	1	RESA_PLAFF	P13830	plasmodium
353	6	1.4	1082	1	SP23_YEAST	P35210	saccharomyc
354	6	1.4	1110	1	TRA1_CAEEL	P34708	caenorhabdi
355	6	1.4	1133	1	VGLM_HANTB	P28728	hantaan vir
356	6	1.4	1133	1	VGLM_SEOU8	P33455	seoul virus
357	6	1.4	1133	1	VGLM_SEOUS	P17880	seoul virus
358	6	1.4	1134	1	VGLM_SEOUR	P28729	seoul virus
359	6	1.4	1140	1	DDB1_CERAE	P33194	cercopithec
360	6	1.4	1140	1	DDB1_HUMAN	Q16531	h dna damag
361	6	1.4	1178	1	YNI7_YEAST	P48231	saccharomyc
362	6	1.4	1208	1	PER_DROYA	Q24767	drosophila
363	6	1.4	1224	1	PER_DROME	P07663	drosophila
364	6	1.4	1248	1	TOPG_SULAC	Q08582	sulfolobus
365	6	1.4	1275	1	TRP_DROME	P19334	drosophila
366	6	1.4	1541	1	MRP2_RAT	Q63120	rattus norv
367	6	1.4	1545	1	MRP2_HUMAN	Q92887	homo sapien
368	6	1.4	1564	1	MRP2_RABIT	Q28689	oryctolagus
369	6	1.4	1839	1	POLR_EPMV	P20126	eggplant mo
370	6	1.4	1956	1	ATX1_PLAFA	Q04956	plasmodium
371	6	1.4	2139	1	CCAC_MOUSE	Q01815	mus musculu
372	6	1.4	2169	1	CCAC_RAT	P22002	rattus norv
373	6	1.4	2183	1	RRPL_MEASA	P35975	measles vir
374	6	1.4	2183	1	RRPL_MEASE	P12576	measles vir
375	6	1.4	2183	1	RRPL_RINDR	P41357	rinderpest
376	6	1.4	2184	1	RRPL_CDVO	P24658	canine dist
377	6	1.4	2201	1	POLG_CXA9	P21404	c genome po
378	6	1.4	2210	1	RRPO_LYCVA	P14240	lymphocytic
379	6	1.4	2236	1	PYR1_DROME	P05990	drosophila
380	6	1.4	2249	1	190K_RICRI	P15921	rickettsia
381	6	1.4	2672	1	GCN1_YEAST	P33892	saccharomyc
382	6	1.4	2895	1	HYDP_DROME	P51592	drosophila
383	6	1.4	3067	1	CA1C_MOUSE	Q60847	mus musculu
384	6	1.4	3144	1	VP13_YEAST	Q07878	saccharomyc
385	6	1.4	4085	1	RRPA_CVH22	Q05002	human coron
386	6	1.4	4466	1	DYHC_ANTCR	P39057	anthocidari
387	6	1.4	4466	1	DYHC_TRIGR	P23098	tripneustes
388	6	1.4	4486	1	DYH9_HUMAN	Q9nyc9	homo sapien
389	5	1.2	13	1	TEMC_RANTE	P56918	rana tempor
390	5	1.2	15	1	FRE2_LITIN	P82022	litoria inf
391	5	1.2	18	1	MCRB_METTE	P22949	methanosarc
392	5	1.2	19	1	UP28_UPEMJ	P82040	uperoleia m
393	5	1.2	20	1	AMP_FUSNU	P81207	fusobacteri
394	5	1.2	21	1	SYA_RAT	P50475	rattus norv
395	5	1.2	21	1	XYNA_DICB4	P80718	dictyoglomu
396	5	1.2	35	1	Y320_BORBU	O51299	borrelia bu
397	5	1.2	36	1	PYY_AMICA	P29205	amia calva
398	5	1.2	37	1	PSBL_ARATH	P29301	arabidopsis

399	5	1.2	37	1	PSBL_ORYSA	P12166	oryza sativ	472	5	1.2	86	1	GON2_CARAU	P51924	carassius a
400	5	1.2	38	1	PSBL_CHLEU	P46306	chlamydomon	473	5	1.2	86	1	OSTZ_YEAST	Q92316	saccharomyc
401	5	1.2	38	1	PSBL_CHLRE	P32974	chlamydomon	474	5	1.2	87	1	AGA2_YEAST	P32781	saccharomyc
402	5	1.2	38	1	PSBL_CHLVU	P56339	chlorella v	475	5	1.2	88	1	ACTO_ACACA	P18281	acanthamoeb
403	5	1.2	38	1	PSBL_CYAPA	P19154	cyanophora	476	5	1.2	88	1	RL15_BRANA	P46289	brassica na
404	5	1.2	38	1	PSBL_EUGGR	P12228	euglena gra	477	5	1.2	88	1	VG86_BPML5	Q05303	mycobacteri
405	5	1.2	38	1	PSBL_GUITH	O78464	guillardia	478	5	1.2	88	1	YAYD_SCHPO	Q10220	schizosacch
406	5	1.2	38	1	PSBL_MARPO	P12165	marchantia	479	5	1.2	89	1	ACBP_GOSHI	Q39779	gossypium h
407	5	1.2	38	1	PSBL_ODOSI	P49514	odontella s	480	5	1.2	89	1	TATA_HAEIN	P57046	haemophilus
408	5	1.2	38	1	PSBL_PINTH	P41617	pinus thunb	481	5	1.2	90	1	DBH_BACST	P02346	bacillus st
409	5	1.2	38	1	PSBL_PORPU	P51389	porphyra pu	482	5	1.2	90	1	MOAD_SYNP7	Q56209	synechococc
410	5	1.2	39	1	VPU_HV1RH	P12518	human immun	483	5	1.2	90	1	RS19_HAEIN	P44385	haemophilus
411	5	1.2	46	1	DIUH_PERAM	P41538	periplaneta	484	5	1.2	90	1	VA14_VACCC	P20991	vaccinia vi
412	5	1.2	48	1	RL32_HELPJ	Q9zmn2	helicobacte	485	5	1.2	90	1	VA14_VARV	P33839	variola vir
413	5	1.2	48	1	RL32_HELPY	P56054	helicobacte	486	5	1.2	91	1	NULM_BRALA	O79420	branchiosto
414	5	1.2	49	1	FER_ACIIN	P81541	acidianus i	487	5	1.2	91	1	RS19_BACST	P12731	bacillus st
415	5	1.2	49	1	Y195_BPT7	P03804	bacteriopha	488	5	1.2	91	1	RS19_BACSU	P21476	bacillus su
416	5	1.2	50	1	YPUE_BACSU	P35153	bacillus su	489	5	1.2	91	1	RS19_ECOLI	P02375	escherichia
417	5	1.2	51	1	IP21_SOLTU	P01079	solanum tub	490	5	1.2	92	1	ACBP_ARATH	P57752	arabidopsis
418	5	1.2	51	1	SURE_HAESO	P36686	haemophilus	491	5	1.2	92	1	DEFI_DROME	P36192	drosophila
419	5	1.2	52	1	ATP8_ALBTU	Q08076	albinaria t	492	5	1.2	92	1	RR19_GUITH	O46898	guillardia
420	5	1.2	52	1	FLMA_ECOLI	P16077	escherichia	493	5	1.2	92	1	RR19_ODOSI	P49506	odontella s
421	5	1.2	52	1	HOK_ECOLI	P11895	escherichia	494	5	1.2	92	1	RR19_PORPU	P51310	porphyra pu
422	5	1.2	53	1	FERB_SULME	P81539	sulfolobus	495	5	1.2	92	1	RS19_BACHD	Q9z910	bacillus ha
423	5	1.2	53	1	RL40_EIMBO	P46575	eimeria bov	496	5	1.2	92	1	RS19_THETH	P80381	thermus aqu
424	5	1.2	54	1	ATP8_CARAU	O78683	carassius a	497	5	1.2	92	1	S11K_ROTPY	Q03056	porcine rot
425	5	1.2	54	1	ATP8_CYPCA	P24948	cyprinus ca	498	5	1.2	92	1	S11K_ROTS1	P11203	simian 11 r
426	5	1.2	54	1	FERA_SULME	P81543	sulfolobus	499	5	1.2	92	1	Y366_METJA	Q57812	methanococc
427	5	1.2	55	1	IP21_CAPAN	P56615	capsicum an	500	5	1.2	93	1	DBH_STRCO	O06447	streptomyce
428	5	1.2	56	1	RL32_ECOLI	P02435	escherichia	501	5	1.2	93	1	RS19_STRPN	Q9ww12	streptococc
429	5	1.2	57	1	V193_BPT7	P03790	bacteriopha	502	5	1.2	93	1	VP44_BPAPS	Q9t1q4	bacteriopha
430	5	1.2	58	1	OM07_CAEEL	P34660	caenorhabdi	503	5	1.2	94	1	IHFB_BUCAI	P57394	buchnera ap
431	5	1.2	59	1	SCK2_ANDMA	P45696	androctonus	504	5	1.2	94	1	YH38_MYCTU	P71996	mycobacteri
432	5	1.2	59	1	SSPF_BACCE	P52970	bacillus ce	505	5	1.2	94	1	YNAI_BACSU	P94487	bacillus su
433	5	1.2	60	1	YOAG_ECOLI	P76247	escherichia	506	5	1.2	95	1	ACP_STRRM	P43677	streptomyce
434	5	1.2	61	1	YFOB_STAEP	Q03382	staphylococ	507	5	1.2	95	1	CILG_HAEIN	P44461	haemophilus
435	5	1.2	62	1	CECB_HYACE	P01508	hyalophora	508	5	1.2	95	1	GON1_PAGMA	P70074	pagrus majo
436	5	1.2	62	1	RS21_ANAVA	P49224	anabaena va	509	5	1.2	95	1	RS19_THEMA	P38520	thermotoga
437	5	1.2	62	1	YCF9_PORPU	P51316	porphyra pu	510	5	1.2	96	1	CY22_RHOFU	P00089	rhodospiril
438	5	1.2	63	1	YDFD_ECOLI	P29010	escherichia	511	5	1.2	96	1	Y039_CHLPN	Q9z9e2	chlamydia p
439	5	1.2	65	1	AMP1_MELGA	P80391	meleagris g	512	5	1.2	96	1	Y335_CHLTR	O84339	chlamydia t
440	5	1.2	65	1	PAFP_PHYAM	P81418	phytolacca	513	5	1.2	96	1	YC4_TYLCV	P27271	tomato yell
441	5	1.2	65	1	VF08_VACCC	P21017	vaccinia vi	514	5	1.2	97	1	CY22_RHOMO	P00088	rhodospiril
442	5	1.2	65	1	VF08_VACCV	P24360	vaccinia vi	515	5	1.2	97	1	VMT2_IAANN	P21430	influenza a
443	5	1.2	65	1	VF08_VARV	P33868	variola vir	516	5	1.2	97	1	YG79_METJA	Q59073	methanococc
444	5	1.2	66	1	HST_YERKR	P31518	yersinia kr	517	5	1.2	97	1	YQIH_BACSU	P40770	bacillus su
445	5	1.2	66	1	RL29_BACST	P04457	bacillus st	518	5	1.2	98	1	SZ10_MOUSE	P17515	mus musculu
446	5	1.2	66	1	RL29_BACSU	P12873	bacillus su	519	5	1.2	98	1	SZ10_RAT	P48973	rattus norv
447	5	1.2	67	1	ATP8_PHOVI	Q00522	phoca vitul	520	5	1.2	99	1	C555_PROAE	P00124	prosthecoch
448	5	1.2	67	1	IBB1_DIOGL	P82469	dioclea gla	521	5	1.2	99	1	FER2_APHSA	P00251	aphanothece
449	5	1.2	67	1	RL29_BACHD	Q9z9k6	bacillus ha	522	5	1.2	99	1	GAS2_ARATH	P46688	arabidopsis
450	5	1.2	67	1	RL30_THEMA	Q9x1j1	thermotoga	523	5	1.2	99	1	GAS3_ARATH	P46687	arabidopsis
451	5	1.2	68	1	ITH5_CUCMA	P19873	cucurbita m	524	5	1.2	99	1	KAP3_MOUSE	P31324	mus musculu
452	5	1.2	68	1	LPPL_MOUSE	P97400	mus musculu	525	5	1.2	100	1	2SS4_CAPMA	P80353	capparis ma
453	5	1.2	70	1	RPBX_YEAST	P22139	saccharomyc	526	5	1.2	100	1	APA2_HUMAN	P02652	homo sapien
454	5	1.2	71	1	ARRD_RAT	P36577	rattus norv	527	5	1.2	100	1	APA2_MACFA	P18656	macaca fasc
455	5	1.2	71	1	YRL5_MYCCA	P43040	mycoplasma	528	5	1.2	100	1	URE3_BORBR	O06705	bordetella
456	5	1.2	72	1	YEJL_HAEIN	P44897	haemophilus	529	5	1.2	100	1	URE3_SYNY3	P73796	synechocyst
457	5	1.2	72	1	YIIU_HAEIN	P44812	haemophilus	530	5	1.2	100	1	VG10_HSVEB	P28980	equine herp
458	5	1.2	73	1	MPT5_AMBTR	P10414	ambrosia tr	531	5	1.2	101	1	YIK2_YEAST	P40488	saccharomyc
459	5	1.2	73	1	RS19_YERPS	P11256	yersinia ps	532	5	1.2	102	1	108_LYCES	Q43495	lycopersico
460	5	1.2	74	1	BR2F_RANES	P40842	rana escule	533	5	1.2	102	1	GP4D_CHLPS	Q46262	chlamydia p
461	5	1.2	74	1	R17E_AERPE	Q9ya67	aeropyrum p	534	5	1.2	102	1	URE2_BORBR	O06707	bordetella
462	5	1.2	75	1	MYR2_MYRPI	Q26464	myrmecia pi	535	5	1.2	102	1	Y116_SSV1	P20207	sulfolobus
463	5	1.2	76	1	ACP_ANAVA	P20803	anabaena va	536	5	1.2	103	1	FER_ACIAM	P49949	acidianus a
464	5	1.2	76	1	ACP_HAEIN	P43709	haemophilus	537	5	1.2	103	1	FER_SULAC	P00219	sulfolobus
465	5	1.2	77	1	ACP_SYNY3	P20804	synechocyst	538	5	1.2	103	1	FER_SULS7	P55907	sulfolobus
466	5	1.2	78	1	IMMV_ECOLI	P22521	escherichia	539	5	1.2	103	1	Y4EI_RHISN	P55432	rhizobium s
467	5	1.2	79	1	GP56_BPSP1	O48410	bacteriopha	540	5	1.2	104	1	2SS1_CAPMA	P80351	capparis ma
468	5	1.2	80	1	GLPF_STRCO	P19255	streptomyce	541	5	1.2	104	1	2SS3_CAPMA	P80352	capparis ma
469	5	1.2	82	1	CYTA_HELAN	Q10992	helianthus	542	5	1.2	104	1	Y338_METJA	Q57784	methanococc
470	5	1.2	82	1	V082_ASFB7	P18557	african swi	543	5	1.2	105	1	COTW_BACSU	Q08310	bacillus su
471	5	1.2	83	1	DCOG_KLEPN	P13155	klebsiella	544	5	1.2	105	1	YB05_FOWPM	P14363	fowlpox vir

545	5	1.2	105	1	YD15_HAEIN	P71375	haemophilus
546	5	1.2	105	1	YIR5_YEAST	P40436	saccharomyc
547	5	1.2	105	1	YM74_MYCTU	Q50689	mycobacteri
548	5	1.2	106	1	ILB1_CAEEL	Q09626	caenorhabdi
549	5	1.2	106	1	Y402_METJA	Q57845	methanococc
550	5	1.2	106	1	YCDZ_SALTY	O54290	salmonella
551	5	1.2	106	1	YPC4_CLAPU	P22369	claviceps p
552	5	1.2	107	1	CPCF_MASLA	P29734	mastigoclad
553	5	1.2	107	1	CYT2_ORYSA	P20907	oryza sativ
554	5	1.2	107	1	FTSL_HAEIN	P45058	haemophilus
555	5	1.2	107	1	YS1_FOAMV	P14352	human spuma
556	5	1.2	108	1	CYC2_ASCSU	P92505	ascaris suu
557	5	1.2	108	1	DBH_BORPR	Q45352	borrelia pa
558	5	1.2	108	1	DBH_BORTU	Q45722	borrelia tu
559	5	1.2	108	1	KV3V_MOUSE	P01674	mus musculu
560	5	1.2	109	1	GLR1_SYNY3	P74593	synechocyst
561	5	1.2	109	1	R14H_BACSH	P39044	bacillus sp
562	5	1.2	109	1	THI2_SYNY3	P73263	synechocyst
563	5	1.2	110	1	2SSI_BRANA	P24565	brassica na
564	5	1.2	110	1	NOLE_RHILP	P23716	rhizobium l
565	5	1.2	111	1	CB2I_CARPI	P38582	carnobacter
566	5	1.2	111	1	LECS_RANJA	P18839	rana japoni
567	5	1.2	111	1	NODC_RHITR	Q53513	rhizobium t
568	5	1.2	111	1	YGN2_YEAST	P53126	saccharomyc
569	5	1.2	111	1	YJC8_YEAST	P47062	saccharomyc
570	5	1.2	112	1	GLNB_PORPU	P51254	porphyra pu
571	5	1.2	112	1	GLNB_RHIET	O54053	rhizobium e
572	5	1.2	112	1	MYR1_MYRPI	Q07932	myrmecia pi
573	5	1.2	112	1	V42_BPT7	P03783	bacteriopha
574	5	1.2	112	1	YC45_METJA	Q58642	methanococc
575	5	1.2	113	1	DAD1_PIG	Q29036	sus scrofa
576	5	1.2	113	1	HEMT_SIPCU	P22766	siphonosoma
577	5	1.2	113	1	HV47_MOUSE	P01823	mus musculu
578	5	1.2	113	1	MEA1_PIG	Q95313	sus scrofa
579	5	1.2	113	1	PRVB_GADCA	P02622	gadus calla
580	5	1.2	113	1	RL24_MICLU	P33103	micrococcus
581	5	1.2	113	1	YNFB_ECOLI	P76170	escherichia
582	5	1.2	114	1	YHT8_YEAST	P38841	saccharomyc
583	5	1.2	115	1	HV44_MOUSE	P01820	mus musculu
584	5	1.2	115	1	NU3M_SHEEP	O78753	ovis aries
585	5	1.2	115	1	PRSD_ECOLI	P42183	escherichia
586	5	1.2	115	1	Y4HO_RHISN	P50359	rhizobium s
587	5	1.2	115	1	YF36_HAEIN	P44247	haemophilus
588	5	1.2	115	1	YT3D_STRCO	Q9xai3	streptomyce
589	5	1.2	115	1	YV1_TYLCM	P27270	tomato yell
590	5	1.2	116	1	MCS_HUMAN	P49901	homo sapien
591	5	1.2	116	1	NU3M_MYTED	Q00568	mytilus edu
592	5	1.2	116	1	NU3M_ONCKI	P20687	oncorhynchu
593	5	1.2	116	1	NU3M_ONCMA	Q37108	oncorhynchu
594	5	1.2	116	1	Y960_HAEIN	P44084	haemophilus
595	5	1.2	117	1	EBRB_BACSU	O31791	bacillus su
596	5	1.2	117	1	GEF2_HUMAN	O08765	homo sapien
597	5	1.2	117	1	HV02_CANFA	P01785	canis famil
598	5	1.2	117	1	SPT4_HUMAN	Q16550	homo sapien
599	5	1.2	117	1	YRGK_CAEEL	Q93408	caenorhabdi
600	5	1.2	118	1	MAR1_HUMAN	Q16655	homo sapien
601	5	1.2	118	1	YC87_MYCPN	P75490	mycoplasma
602	5	1.2	118	1	YG35_BPLLH	Q04769	lactococcus
603	5	1.2	118	1	YMX2_YEAST	Q04276	saccharomyc
604	5	1.2	119	1	EYA3_CHICK	Q9yh99	gallus gall
605	5	1.2	119	1	EYA4_CHICK	Q9yh98	gallus gall
606	5	1.2	119	1	EYA4_FUGRU	Q9yha1	fugu rubrip
607	5	1.2	119	1	GLNA_TRITH	P51120	trichodesmi
608	5	1.2	119	1	HIS3_RHOSH	Q53158	rhodobacter
609	5	1.2	119	1	YVAE_BACSU	O32227	bacillus su
610	5	1.2	120	1	C550_BACSU	P24469	bacillus su
611	5	1.2	120	1	CHH1_PENJP	O15980	penaeus jap
612	5	1.2	120	1	CHH2_PENJP	Q9u5d2	penaeus jap
613	5	1.2	120	1	SY16_HUMAN	O15467	homo sapien
614	5	1.2	120	1	YJ9V_YEAST	P47184	saccharomyc
615	5	1.2	121	1	V121_ASFL5	P26705	african swi
616	5	1.2	122	1	CADC_STAAU	P20047	staphylococ
617	5	1.2	122	1	RL14_CHLPN	Q9z7r7	chlamydia p
618	5	1.2	122	1	RL14_CHLTR	P28533	chlamydia t
619	5	1.2	122	1	RL14_THEMA	P38508	thermotoga
620	5	1.2	122	1	YPS2_PLEBO	P15734	plectonema
621	5	1.2	123	1	FUR_PSEFL	O68563	pseudomonas
622	5	1.2	123	1	HEB2_ANAVA	P46052	anabaena va
623	5	1.2	123	1	MIR1_RAT	Q9wtw0	rattus norv
624	5	1.2	123	1	NUOA_RICPR	Q9zdh1	rickettsia
625	5	1.2	123	1	YB8F_YEAST	P38357	saccharomyc
626	5	1.2	123	1	YBAV_ECOLI	P77415	escherichia
627	5	1.2	124	1	YC69_MYCTU	Q11050	mycobacteri
628	5	1.2	125	1	ACPM_YEAST	P32463	saccharomyc
629	5	1.2	125	1	FABL_CHICK	P80226	gallus gall
630	5	1.2	125	1	NEU1_HUMAN	P01178	homo sapien
631	5	1.2	125	1	NEU1_PIG	P01177	sus scrofa
632	5	1.2	125	1	ULA6_HCMVA	P16737	human cytom
633	5	1.2	126	1	CD59_PAPSP	Q28785	papio sp. (
634	5	1.2	126	1	SYGB_NEIGO	Q50945	neisseria g
635	5	1.2	126	1	YEA8_YEAST	P40001	saccharomyc
636	5	1.2	126	1	YF83_METJA	Q58978	methanococc
637	5	1.2	127	1	ALL1_SINAL	P15322	sinapis alb
638	5	1.2	127	1	PFD6_MOUSE	Q03958	mus musculu
639	5	1.2	127	1	SAA_ANAPL	P02740	anas platyr
640	5	1.2	127	1	Y55B_MYCPN	P75047	mycoplasma
641	5	1.2	127	1	YKH6_YEAST	P36082	saccharomyc
642	5	1.2	128	1	CD59_AOTTR	P51447	aotus trivi
643	5	1.2	128	1	CD59_CALSQ	P46657	callithrix
644	5	1.2	128	1	CD59_HUMAN	P13987	homo sapien
645	5	1.2	128	1	VNBP_SHVX	Q04580	shallot vir
646	5	1.2	129	1	ALL1_BRAJU	P80207	brassica ju
647	5	1.2	129	1	HV2F_HUMAN	P01824	homo sapien
648	5	1.2	129	1	MINK_FELCA	Q9xspl	felis silve
649	5	1.2	129	1	PFD6_HUMAN	O15212	homo sapien
650	5	1.2	129	1	RS17_CHICK	P08636	gallus gall
651	5	1.2	129	1	YKR4_CAEEL	P34310	caenorhabdi
652	5	1.2	130	1	ITRY_SINAR	P38057	sinapis arv
653	5	1.2	130	1	NU5M_STRFN	Q35832	strongyloce
654	5	1.2	131	1	RS8_BACSU	P12879	bacillus su
655	5	1.2	131	1	TAP2_ANTMA	P24609	antirrhinum
656	5	1.2	132	1	ANFG_AZOPA	O68943	azotobacter
657	5	1.2	132	1	H2A1_LEIDO	P27891	leishmania
658	5	1.2	132	1	H2A2_LEIDO	P27892	leishmania
659	5	1.2	132	1	HBAM_RANCA	P02022	rana catesb
660	5	1.2	132	1	MRKF_KLEPN	P21650	klebsiella
661	5	1.2	132	1	RS19_PYRHO	O59422	pyrococcus
662	5	1.2	132	1	RS8_MYCTU	P95066	mycobacteri
663	5	1.2	132	1	VPF1_LAMBD	P03709	bacteriopha
664	5	1.2	132	1	YDDE_SALTY	P40788	salmonella
665	5	1.2	133	1	RBFA_SYNY3	Q55625	synechocyst
666	5	1.2	133	1	RL22_AQUAE	O66436	aquifex aeo
667	5	1.2	133	1	RL22_AQUPY	Q9zi45	aquifex pyr
668	5	1.2	133	1	RS8_METTH	O26126	methanobact
669	5	1.2	134	1	FUR_PSEAE	Q03456	pseudomonas
670	5	1.2	134	1	FUR_PSEPU	Q52083	pseudomonas
671	5	1.2	134	1	IL5_MACMU	P48093	macaca mula
672	5	1.2	134	1	RS17_CRIGR	P06584	cricetulus
673	5	1.2	134	1	RS17_HUMAN	P08708	homo sapien
674	5	1.2	134	1	SY21_HUMAN	O00585	homo sapien
675	5	1.2	134	1	VA2_BPT5	P23541	bacteriopha
676	5	1.2	134	1	YCGX_ECOLI	P75988	escherichia
677	5	1.2	135	1	CHH_MACLE	O77220	macrobrachi
678	5	1.2	135	1	CYT1_MAIZE	P31726	zea mays (m
679	5	1.2	135	1	HNS_BUCAI	P57360	buchnera ap
680	5	1.2	135	1	LECG_CROAT	P21963	crotalus at
681	5	1.2	135	1	Y548_BUCAI	P57613	buchnera ap
682	5	1.2	136	1	RL16_ECOLI	P02414	escherichia
683	5	1.2	136	1	Y354_MYCPN	P75248	mycoplasma
684	5	1.2	136	1	Y756_METJA	Q58166	methanococc
685	5	1.2	137	1	UK14_CAPHI	P80601	capra hircu
686	5	1.2	137	1	UK14_HUMAN	P52758	homo sapien
687	5	1.2	137	1	YCDP_ECOLI	P75904	escherichia
688	5	1.2	137	1	YUXK_BACSU	P40761	bacillus su
689	5	1.2	138	1	ATPE_BUCAI	P57125	buchnera ap
690	5	1.2	138	1	BGAL_AGRRD	P30812	agrobacteri

691	5	1.2	138	1	LSHB_MACRU	O46483	macropus ru
692	5	1.2	138	1	POLG_HCVE0	P27953	hepatitis c
693	5	1.2	138	1	VA23_BPBF2	P19348	bacteriopha
694	5	1.2	138	1	YIGG_ECOLI	P27843	escherichia
695	5	1.2	139	1	FXSA_SERMA	P37148	serratia ma
696	5	1.2	139	1	H163_HUMAN	Q9p003	homo sapien
697	5	1.2	139	1	MTAB_SYNP2	P34883	synechococc
698	5	1.2	139	1	RS12_MYCGE	P47333	mycoplasma
699	5	1.2	139	1	RS12_MYCPN	P75546	mycoplasma
700	5	1.2	139	1	VG08_BPMU	Q38480	bacteriopha
701	5	1.2	139	1	VP26_HSVSA	Q01047	herpesvirus
702	5	1.2	139	1	Y4VR_RHISN	P55677	rhizobium s
703	5	1.2	139	1	YE8A_METJA	P81331	methanococc
704	5	1.2	140	1	YPHA_ECOLI	P77751	escherichia
705	5	1.2	141	1	HBA1_TORMA	P20244	torpedo mar
706	5	1.2	141	1	HBA2_TORMA	P20245	torpedo mar
707	5	1.2	141	1	HBA2_VAREX	P18981	varanus exa
708	5	1.2	141	1	LSHB_TRIVU	O46482	trichosurus
709	5	1.2	141	1	RS8_MYCGE	P47411	mycoplasma
710	5	1.2	141	1	YPRO_OWEFU	P21260	owenia fusi
711	5	1.2	142	1	TCA_HUMAN	P01848	homo sapien
712	5	1.2	143	1	LGB1_SOYBN	P02235	glycine max
713	5	1.2	143	1	LGBA_SOYBN	P02238	glycine max
714	5	1.2	143	1	PBP5_DROME	P54195	drosophila
715	5	1.2	143	1	PER_DROPI	Q25206	drosophila
716	5	1.2	144	1	HV43_MOUSE	P01819	mus musculu
717	5	1.2	144	1	LGB3_SOYBN	P02237	glycine max
718	5	1.2	144	1	YAT3_STAAU	P52080	staphylococ
719	5	1.2	144	1	YIT5_YEAST	P40571	saccharomyc
720	5	1.2	145	1	ANG3_MOUSE	P97802	mus musculu
721	5	1.2	145	1	GLB2_LUMTE	P02218	lumbricus t
722	5	1.2	145	1	HIT_BACSU	O07513	bacillus su
723	5	1.2	145	1	LGBA_PHAVU	P02234	phaseolus v
724	5	1.2	145	1	LGB_PSOTE	P27199	psophocarpu
725	5	1.2	145	1	NUIM_TRYBB	P30826	trypanosoma
726	5	1.2	145	1	PA20_BUNMU	P00606	bungarus mu
727	5	1.2	145	1	RL32_AERPE	Q9yf92	aeropyrum p
728	5	1.2	145	1	SODM_STRIN	O54233	streptococc
729	5	1.2	145	1	Y4WH_RHISN	P55686	rhizobium s
730	5	1.2	145	1	YGL8_YEAST	P53132	saccharomyc
731	5	1.2	145	1	YPI2_AGRTU	Q44433	agrobacteri
732	5	1.2	146	1	GLB_DOLAU	P09965	dolabella a
733	5	1.2	146	1	HBB1_IGUIG	P18987	iguana igua
734	5	1.2	146	1	HBB1_VAREX	P18993	varanus exa
735	5	1.2	146	1	HBB1_XENTR	P07429	xenopus tro
736	5	1.2	146	1	HBB2_NAJNA	P22743	naja naja (
737	5	1.2	146	1	HBB_ACCGE	P08851	accipiter g
738	5	1.2	146	1	HBB_AEGMO	P07418	aegypius mo
739	5	1.2	146	1	HBB_ANAPL	P02114	anas platyr
740	5	1.2	146	1	HBB_ANAPP	P02115	anas platyr
741	5	1.2	146	1	HBB_ANSAN	P02117	anser anser
742	5	1.2	146	1	HBB_ANSIN	P02118	anser indic
743	5	1.2	146	1	HBB_ANSSE	P02120	anseranas s
744	5	1.2	146	1	HBB_APTFO	P80216	aptenodytes
745	5	1.2	146	1	HBB_APUAP	P15165	apus apus (
746	5	1.2	146	1	HBB_AQUCH	P02122	aquila chry
747	5	1.2	146	1	HBB_ARAAR	P02116	ara araraun
748	5	1.2	146	1	HBB_BRACA	P02119	branta cana
749	5	1.2	146	1	HBB_BRATR	P14526	bradypus tr
750	5	1.2	146	1	HBB_CAIMO	P14260	cairina mos
751	5	1.2	146	1	HBB_CATMA	P82113	catharacta
752	5	1.2	146	1	HBB_CHEKU	P80271	chelodonich
753	5	1.2	146	1	HBB_CHICK	P02112	gallus gall
754	5	1.2	146	1	HBB_CHLME	P07036	chloephaga
755	5	1.2	146	1	HBB_CHRPI	P13274	chrysemys p
756	5	1.2	146	1	HBB_CICCI	P02125	ciconia cic
757	5	1.2	146	1	HBB_COLLI	P11342	columba liv
758	5	1.2	146	1	HBB_COTJA	P30893	coturnix co
759	5	1.2	146	1	HBB_CTEGU	P20855	ctenodactyl
760	5	1.2	146	1	HBB_DASNO	P02087	dasypus nov
761	5	1.2	146	1	HBB_DIDMA	P02109	didelphis m
762	5	1.2	146	1	HBB_ELEMA	P02084	elephas max
763	5	1.2	146	1	HBB_EUDCR	P10058	eudyptes cr
764	5	1.2	146	1	HBB_FRAPO	P22742	francolinus
765	5	1.2	146	1	HBB_LARRI	P08261	larus ridib
766	5	1.2	146	1	HBB_LOXAF	P02085	loxodonta a
767	5	1.2	146	1	HBB_MICGA	P41332	microcephal
768	5	1.2	146	1	HBB_ORNAN	P02111	ornithorhyn
769	5	1.2	146	1	HBB_PASMO	P07406	passer mont
770	5	1.2	146	1	HBB_PHACA	P10782	phalacrocor
771	5	1.2	146	1	HBB_PHACO	P02113	phasianus c
772	5	1.2	146	1	HBB_PHORU	P02121	phoenicopte
773	5	1.2	146	1	HBB_PSIKR	P21668	psittacula
774	5	1.2	146	1	HBB_RHEAM	P02124	rhea americ
775	5	1.2	146	1	HBB_STRCA	P02123	struthio ca
776	5	1.2	146	1	HBB_STUVU	P02126	sturnus vul
777	5	1.2	146	1	HBB_TACAC	P02110	tachyglossu
778	5	1.2	146	1	HBB_TUPGL	P02052	tupaia glis
779	5	1.2	146	1	HBB_TURME	P14524	turdus meru
780	5	1.2	146	1	HBB_VULGR	P07411	vultur gryp
781	5	1.2	146	1	HBE1_CAPHI	P02102	capra hircu
782	5	1.2	146	1	HBE_AOTAZ	P43350	aotus azara
783	5	1.2	146	1	HBE_ATEBE	P51438	ateles belz
784	5	1.2	146	1	HBE_CALJA	P51439	callithrix
785	5	1.2	146	1	HBE_CEBAL	Q28220	cebus albif
786	5	1.2	146	1	HBE_CHEME	Q28338	cheirogaleu
787	5	1.2	146	1	HBE_DAUMA	Q28356	daubentonia
788	5	1.2	146	1	HBE_DIDMA	P11025	didelphis m
789	5	1.2	146	1	HBE_EULFU	P08223	eulemur ful
790	5	1.2	146	1	HBE_GALCR	P19759	galago cras
791	5	1.2	146	1	HBE_HUMAN	P02100	homo sapien
792	5	1.2	146	1	HBE_HYLSY	Q95190	hylobates s
793	5	1.2	146	1	HBE_LAGLA	P51440	lagothrix l
794	5	1.2	146	1	HBE_LEORO	P51441	leontopithe
795	5	1.2	146	1	HBE_MACMU	Q28507	macaca mula
796	5	1.2	146	1	HBE_MICMU	Q28496	microcebus
797	5	1.2	146	1	HBE_PANPA	Q28775	pan paniscu
798	5	1.2	146	1	HBE_PITIR	P51442	pithecia ir
799	5	1.2	146	1	HBE_PONPY	P18994	pongo pygma
800	5	1.2	146	1	HBE_PROVE	Q95238	propithecus
801	5	1.2	146	1	HBE_RABIT	P02103	oryctolagus
802	5	1.2	146	1	HBE_SAGMI	P51443	saguinus mi
803	5	1.2	146	1	HBE_SAISC	P51444	saimiri sci
804	5	1.2	146	1	HBE_SMICR	Q28931	sminthopsis
805	5	1.2	146	1	HBE_TARSY	P18435	tarsius syr
806	5	1.2	146	1	HBT_PIG	P04246	sus scrofa
807	5	1.2	146	1	RM27_YEAST	P36526	saccharomyc
808	5	1.2	146	1	RS13_ARCFU	O27999	archaeoglob
809	5	1.2	146	1	SMD1_YEAST	Q02260	saccharomyc
810	5	1.2	146	1	YAAR_BACSU	P37539	bacillus su
811	5	1.2	147	1	GRP_RAT	P24393	rattus norv
812	5	1.2	147	1	IP2T_SOLTU	Q41435	solanum tub
813	5	1.2	147	1	IP2Y_SOLTU	Q41489	solanum tub
814	5	1.2	147	1	YDAA_SCHPO	Q10352	schizosacch
815	5	1.2	147	1	YDEF_SCHPO	Q10448	schizosacch
816	5	1.2	147	1	YGY0_YEAST	P53074	saccharomyc
817	5	1.2	147	1	YLYA_BACSU	Q45478	bacillus su
818	5	1.2	147	1	YPMF_ECOLI	P07116	escherichia
819	5	1.2	148	1	COX6_YEAST	P00427	saccharomyc
820	5	1.2	148	1	IP21_LYCES	P05119	lycopersico
821	5	1.2	148	1	NUSB_AQUAE	O66530	aquifex aeo
822	5	1.2	148	1	RS13_PYRHO	O74021	pyrococcus
823	5	1.2	148	1	SSB_HUMAN	Q04837	homo sapien
824	5	1.2	148	1	VFUS_CAPVK	P16717	capripoxvir
825	5	1.2	148	1	YE84_HELPJ	Q9zjd5	helicobacte
826	5	1.2	148	1	YFDP_ECOLI	P76512	escherichia
827	5	1.2	149	1	RL13_SULSO	P95991	sulfolobus
828	5	1.2	149	1	YB07_METJA	Q58507	methanococc
829	5	1.2	150	1	CALM_SCHPO	P05933	schizosacch
830	5	1.2	150	1	PBP2_DROME	P54192	drosophila
831	5	1.2	150	1	Y404_METJA	Q57847	methanococc
832	5	1.2	150	1	YBCV_ECOLI	P77598	escherichia
833	5	1.2	150	1	YCCA_BACP3	P55815	bacillus ps
834	5	1.2	150	1	YM97_MYCTU	Q50669	mycobacteri
835	5	1.2	150	1	YTR1_AZOBR	P26943	azospirillu
836	5	1.2	151	1	CSGB_SALTY	P55226	salmonella

837	5	1.2	151	1	EP1_HUMAN	Q15668	homo sapien
838	5	1.2	151	1	MLE_HALRO	P07462	halocynthia
839	5	1.2	151	1	SSB_RAT	P28042	rattus norv
840	5	1.2	151	1	VG10_HSVI1	Q00134	ictalurid h
841	5	1.2	151	1	Y903_AQUAE	O67050	aquifex aeo
842	5	1.2	151	1	YDS8_SCHPO	O14183	schizosacch
843	5	1.2	151	1	YKA2_YEAST	P36108	saccharomyc
844	5	1.2	152	1	RS15_ARATH	Q08112	arabidopsis
845	5	1.2	152	1	RS15_PODAN	P34737	podospora a
846	5	1.2	152	1	RS18_ARATH	P34788	arabidopsis
847	5	1.2	152	1	SODC_NPVOP	O12933	orgyia pseu
848	5	1.2	153	1	FLAP_PIG	P30356	sus scrofa
849	5	1.2	153	1	HEMA_MUMPS	P33480	mumps virus
850	5	1.2	153	1	HS11_SOYBN	P02519	glycine max
851	5	1.2	153	1	HS13_SOYBN	P04793	glycine max
852	5	1.2	153	1	IAAT_AZOBR	P26945	azospirillu
853	5	1.2	153	1	IP2K_SOLTU	P01080	solanum tub
854	5	1.2	153	1	MPP3_BOVIN	P80195	bos taurus
855	5	1.2	153	1	MYG_DIDMA	P02193	didelphis m
856	5	1.2	153	1	MYG_ELEMA	P02186	elephas max
857	5	1.2	153	1	MYG_ERIEU	P02156	erinaceus e
858	5	1.2	153	1	MYG_HUMAN	P02144	homo sapien
859	5	1.2	153	1	MYG_LOXAF	P02187	loxodonta a
860	5	1.2	153	1	MYG_ORYAF	P02164	orycteropus
861	5	1.2	153	1	MYG_PANTR	P02145	pan troglod
862	5	1.2	153	1	MYG_ROUAE	P02163	rousettus a
863	5	1.2	153	1	YORZ_LISMO	P33385	listeria mo
864	5	1.2	153	1	YPH1_SYNP2	P32039	synechococc
865	5	1.2	153	1	YSI2_CAEEL	Q10034	caenorhabdi
866	5	1.2	153	1	YZ30_METJA	Q60288	methanococc
867	5	1.2	154	1	HS14_SOYBN	P04794	glycine max
868	5	1.2	154	1	HS15_SOYBN	P04795	glycine max
869	5	1.2	154	1	IP25_SOLTU	Q41488	solanum tub
870	5	1.2	154	1	IP27_SOLTU	Q43652	solanum tub
871	5	1.2	154	1	KMLS_MELGA	P56276	meleagris g
872	5	1.2	154	1	PTGA_MYCCA	P45618	mycoplasma
873	5	1.2	155	1	PRS1_SOLTU	P17641	solanum tub
874	5	1.2	155	1	PRS2_SOLTU	P17642	solanum tub
875	5	1.2	155	1	YBEA_ECOLI	P05850	escherichia
876	5	1.2	155	1	YBEA_HAEIN	P44470	haemophilus
877	5	1.2	156	1	BCCP_PSEAE	P37799	pseudomonas
878	5	1.2	156	1	ISPF_AQUAE	O67089	aquifex aeo
879	5	1.2	156	1	YLXS_BACSU	P32726	bacillus su
880	5	1.2	157	1	FLIN_PSEAE	Q51466	pseudomonas
881	5	1.2	157	1	SSSP_PLAFO	P17503	plasmodium
882	5	1.2	157	1	UBC3_SCHPO	P40984	schizosacch
883	5	1.2	158	1	IP2X_SOLTU	Q00782	solanum tub
884	5	1.2	158	1	LECG_TRIST	Q9ygp1	trimeresuru
885	5	1.2	158	1	YV02_HUMAN	Q99765	homo sapien
886	5	1.2	159	1	NIFZ_AZOVI	P14889	azotobacter
887	5	1.2	159	1	UL14_PRVN3	P30661	pseudorabie
888	5	1.2	159	1	YICN_ECOLI	P31439	escherichia
889	5	1.2	160	1	CRTK_RHOCA	P17057	rhodobacter
890	5	1.2	160	1	MLR_SPISA	P08051	spisula sac
891	5	1.2	160	1	PHAA_SYNP6	P06112	synechococc
892	5	1.2	161	1	HS16_SOYBN	P05478	glycine max
893	5	1.2	161	1	LINA_STAHA	P06107	staphylococ
894	5	1.2	161	1	PHAB_AGLNE	P28556	aglaothamni
895	5	1.2	161	1	SIXA_ECOLI	P76502	escherichia
896	5	1.2	161	1	Y311_RICPR	Q9zdl6	rickettsia
897	5	1.2	161	1	YC67_EUGGR	P48337	euglena gra
898	5	1.2	161	1	YEP7_YEAST	P40043	saccharomyc
899	5	1.2	161	1	YLAL_BACSU	O07636	bacillus su
900	5	1.2	162	1	HXA9_CAVPO	P51783	cavia porce
901	5	1.2	163	1	DNE1_CHLRE	P05725	chlamydomon
902	5	1.2	163	1	OB_CHICK	O42164	gallus gall
903	5	1.2	163	1	PTFB_BACSU	P26380	bacillus su
904	5	1.2	163	1	SAK_BPP42	P15240	bacteriopha
905	5	1.2	163	1	SAK_STAAU	P00802	staphylococ
906	5	1.2	164	1	2SS1_ARATH	P15457	arabidopsis
907	5	1.2	164	1	2SS3_ARATH	P15459	arabidopsis
908	5	1.2	164	1	GLB1_CHLEU	Q08753	chlamydomon
909	5	1.2	164	1	SIXA_HAEIN	P44164	haemophilus
910	5	1.2	164	1	TSTP_HSVSA	P18347	herpesvirus
911	5	1.2	164	1	UBC7_CAEEL	P34477	caenorhabdi
912	5	1.2	164	1	Y239_METJA	Q57691	methanococc
913	5	1.2	164	1	Y245_MYCPN	P75430	mycoplasma
914	5	1.2	164	1	YCE1_YEAST	P25571	saccharomyc
915	5	1.2	165	1	OGT_MYCLE	P52982	mycobacteri
916	5	1.2	165	1	PMX2_CHICK	Q90963	gallus gall
917	5	1.2	165	1	RL12_CAEEL	O62290	caenorhabdi
918	5	1.2	165	1	RS5_BORBU	O51448	borrelia bu
919	5	1.2	165	1	Y872_RICPR	Q9zc94	rickettsia
920	5	1.2	166	1	2SS4_ARATH	P15460	arabidopsis
921	5	1.2	166	1	PETP_RHOCA	P31078	rhodobacter
922	5	1.2	166	1	VP06_BPPRD	P27379	bacteriopha
923	5	1.2	166	1	Y516_SYNY3	Q55837	synechocyst
924	5	1.2	167	1	CAT3_HUMAN	O15182	homo sapien
925	5	1.2	167	1	CAT3_MOUSE	O35648	mus musculu
926	5	1.2	167	1	OB_MOUSE	P41160	mus musculu
927	5	1.2	167	1	TCTP_PSEMZ	Q9zrx0	pseudotsuga
928	5	1.2	168	1	LSPA_CHLPN	Q9z817	chlamydia p
929	5	1.2	168	1	PTP_NPVAC	P24656	autographa
930	5	1.2	168	1	RL10_BORBU	O51352	borrelia bu
931	5	1.2	168	1	VA40_VACCC	P21063	vaccinia vi
932	5	1.2	168	1	VA40_VACCV	P24765	vaccinia vi
933	5	1.2	169	1	DYR8_ECOLI	Q57452	escherichia
934	5	1.2	169	1	IPYR_SYNY3	P80507	synechocyst
935	5	1.2	169	1	MEMG_METCA	P11987	methylococc
936	5	1.2	169	1	PHAF_CYAPA	P48087	cyanophora
937	5	1.2	170	1	2SS2_ARATH	P15458	arabidopsis
938	5	1.2	170	1	LSPA_BORBU	O51425	borrelia bu
939	5	1.2	170	1	OPR3_NEIMC	P10171	neisseria m
940	5	1.2	171	1	CP18_RABIT	P25230	oryctolagus
941	5	1.2	171	1	GLB2_CHLEU	P52334	chlamydomon
942	5	1.2	171	1	LYCV_BPMU	Q9tlx2	bacteriopha
943	5	1.2	171	1	NSG2_MOUSE	P47759	mus musculu
944	5	1.2	171	1	TOXS_VIBPA	Q05939	vibrio para
945	5	1.2	171	1	Y52L_SYNY3	Q55911	synechocyst
946	5	1.2	172	1	CH18_DROME	P07184	drosophila
947	5	1.2	172	1	IMDH_METKA	P50100	methanopyru
948	5	1.2	172	1	MEA1_HUMAN	Q16626	homo sapien
949	5	1.2	172	1	NU6M_CYPCA	P24982	cyprinus ca
950	5	1.2	172	1	NUDH_VIBCH	Q9ku53	vibrio chol
951	5	1.2	173	1	HR12_MOUSE	P52760	mus musculu
952	5	1.2	173	1	NU6M_CARAU	O78689	carassius a
953	5	1.2	173	1	NU6M_CROLA	P34196	crossostoma
954	5	1.2	173	1	RT3C_ACTPL	Q04474	actinobacil
955	5	1.2	173	1	TCB1_MOUSE	P01852	mus musculu
956	5	1.2	173	1	TCB2_MOUSE	P01851	mus musculu
957	5	1.2	174	1	FRIL_HUMAN	P02792	homo sapien
958	5	1.2	174	1	FRIL_RABIT	P09451	oryctolagus
959	5	1.2	174	1	MEA1_BOVIN	Q29407	bos taurus
960	5	1.2	174	1	MEA1_MOUSE	Q64327	mus musculu
961	5	1.2	175	1	IPYR_ECOLI	P17288	escherichia
962	5	1.2	175	1	MFA3_BOVIN	Q28103	bos taurus
963	5	1.2	175	1	Y901_MYCTU	Q10559	mycobacteri
964	5	1.2	176	1	IL7_BOVIN	P26895	bos taurus
965	5	1.2	176	1	IL7_SHEEP	Q28540	ovis aries
966	5	1.2	176	1	MPV1_HUMAN	P39210	homo sapien
967	5	1.2	176	1	MPV1_MOUSE	P19258	mus musculu
968	5	1.2	176	1	NUSG_HELPJ	Q9zk20	helicobacte
969	5	1.2	176	1	NUSG_HELPY	P55976	helicobacte
970	5	1.2	176	1	V6B_CVFE3	P33467	feline ente
971	5	1.2	177	1	LYCV_BPH19	Q9zxb7	bacteriopha
972	5	1.2	177	1	RL6_BACST	P02391	bacillus st
973	5	1.2	177	1	VN02_VARV	P34019	variola vir
974	5	1.2	178	1	2SS2_BRANA	P01090	brassica na
975	5	1.2	178	1	NI2M_BOVIN	Q02369	bos taurus
976	5	1.2	178	1	NI2M_HUMAN	Q9y6m9	homo sapien
977	5	1.2	178	1	Y057_MYCGE	P47303	mycoplasma
978	5	1.2	178	1	Y4YJ_RHIFR	P72273	rhizobium f
979	5	1.2	178	1	Y4YJ_RHISN	P55718	rhizobium s
980	5	1.2	178	1	YREP_BUCTC	O31285	buchnera ap
981	5	1.2	179	1	ADAB_BACSU	P19220	bacillus su
982	5	1.2	179	1	APT_HELPJ	Q9zlq9	helicobacte

983	5	1.2	179	1	APT_HELPY	O25296	helicobacte
984	5	1.2	179	1	APT_HUMAN	P07741	homo sapien
985	5	1.2	179	1	HPRT_HAEIN	P45078	haemophilus
986	5	1.2	179	1	US25_HCMVA	P17145	human cytom
987	5	1.2	179	1	YAJ3_YEAST	P39546	saccharomyc
988	5	1.2	179	1	YC21_ANTSP	Q06463	antithamnio
989	5	1.2	179	1	YCDZ_ECOLI	P75916	escherichia
990	5	1.2	180	1	APT_CRILO	P47952	cricetulus
991	5	1.2	180	1	APT_GERCA	Q64414	gerbillus c
992	5	1.2	180	1	APT_MASHI	Q64427	mastomys hi
993	5	1.2	180	1	APT_MOUSE	P08030	mus musculu
994	5	1.2	180	1	APT_MUSPA	P47956	mus pahari
995	5	1.2	180	1	APT_MUSSI	P47957	mus spicile
996	5	1.2	180	1	APT_RAT	P36972	rattus norv
997	5	1.2	180	1	APT_STOLO	P47958	stochomys l
998	5	1.2	180	1	ATPD_SYNP6	P08448	synechococc
999	5	1.2	180	1	FMF3_ECOLI	P11312	escherichia
1000	5	1.2	180	1	GCH1_HELPJ	Q9zks2	helicobacte

ALIGNMENTS

```
RESULT   1
PRTP_HSVB2
ID   PRTP_HSVB2     STANDARD;      PRT;   664 AA.
AC   P17587;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROBABLE PROCESSING AND TRANSPORT PROTEIN (INFECTED CELL PROTEIN 18.5)
DE   (ICP18.5 PROTEIN) (72 KDA PROTEIN).
GN   UL28.
OS   Bovine herpesvirus type 2 (strain BMV) (Bovine mammillitis virus).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
OX   NCBI_TaxID=10296;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88306231; PubMed=2841793;
RA   Hammerschmidt W., Conraths F., Mankertz J., Pauli G., Ludwig H.,
RA   Buhk H.-J.;
RT   "Conservation of a gene cluster including glycoprotein B in bovine
RT   herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1).";
RL   Virology 165:388-405(1988).
CC   -!- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLOCATION OF THE VIRUS
CC       GLYCOPROTEINS TO MEMBRANES. IT IS INVOLVED IN CAPSID MATURATION.
CC   -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES PRTP FAMILY.
DR   PIR; B29242; WMBEBH.
DR   InterPro; IPR000501; -.
DR   Pfam; PF01366; PRTP; 1.
KW   Late protein; Capsid assembly.
SQ   SEQUENCE   664 AA;  72367 MW;  A84F7ED8EBE24D29 CRC64;


  Query Match          1.9%;  Score 8;  DB 1;  Length 664;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     624 LKEELARF 631


RESULT   2
PRTP_PRVIF
ID   PRTP_PRVIF     STANDARD;      PRT;   724 AA.
AC   P11871;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROBABLE PROCESSING AND TRANSPORT PROTEIN (INFECTED CELL PROTEIN
DE   18.5).
GN   ICP18.5.
OS   Pseudorabies virus (strain Indiana-Funkhauser / Becker) (PRV).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=31523;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89263808; PubMed=2542904;
RA   Pederson N.E., Enquist L.W.;
RT   "The nucleotide sequence of a pseudorabies virus gene similar to
RT   ICP18.5 of herpes simplex virus type 1.";
RL   Nucleic Acids Res. 17:3597-3597(1989).
CC   -!- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLOCATION OF THE VIRUS
CC       GLYCOPROTEINS TO MEMBRANES. IT IS INVOLVED IN CAPSID MATURATION.
CC   -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES PRTP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X14573; CAA32712.1; -.
DR   PIR; A33779; WMBEPR.
DR   HSSP; P04002; 1WFA.
DR   InterPro; IPR000501; -.
DR   Pfam; PF01366; PRTP; 1.
KW   Late protein; Capsid assembly.
SQ   SEQUENCE   724 AA;  78881 MW;  87C16178DC5BADB4 CRC64;


  Query Match          1.9%;  Score 8;  DB 1;  Length 724;
  Best Local Similarity   100.0%;  Pred. No. 2.7;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     602 LKEELARF 609


RESULT   3
PRTP_HSVEB
ID   PRTP_HSVEB     STANDARD;      PRT;   766 AA.
AC   P28973; P33549;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROBABLE PROCESSING AND TRANSPORT PROTEIN.
GN   32.
OS   Equine herpesvirus type 1 (strain Ab4p) (EHV-1), and
OS   Equine herpesvirus type 1 (isolate HVS25A) (EHV-1).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=31520, 10327;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB4P;
RX   MEDLINE=92295566; PubMed=1318606;
RA   Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT   "The DNA sequence of equine herpesvirus-1.";
RL   Virology 189:304-316(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ISOLATE HVS25A;
RX   MEDLINE=94106109; PubMed=8279122;
RA   Bell C.W., Whalley J.M.;
RT   "Herpesvirus ICP18.5 and DNA-binding protein genes are conserved in
RT   equine herpesvirus-1.";
RL   Virus Genes 7:219-228(1993).
CC   -!- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLOCATION OF THE VIRUS
CC       GLYCOPROTEINS TO MEMBRANES. IT IS INVOLVED IN CAPSID MATURATION.
```

```
CC   -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES PRTP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M86664; AAB02467.1; -.
DR   EMBL; D13930; BAA03034.1; -.
DR   PIR; F36798; WZBEC5.
DR   PIR; JQ0847; WZBEJ5.
DR   InterPro; IPR000501; -.
DR   Pfam; PF01366; PRTP; 1.
KW   Capsid assembly.
SQ   SEQUENCE   766 AA;  84383 MW;  D5B501A161ACCE03 CRC64;


  Query Match             1.9%;  Score 8;  DB 1;  Length 766;
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     629 LKEELARF 636


RESULT   4
PRTP_HSV1F
ID   PRTP_HSV1F     STANDARD;      PRT;   780 AA.
AC   P06490;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROCESSING AND TRANSPORT PROTEIN (INFECTED CELL PROTEIN 18.5) (ICP18.5
DE   PROTEIN).
GN   UL28.
OS   Herpes simplex virus (type 1 / strain F).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
OX   NCBI_TaxID=10304;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87061190; PubMed=3023664;
RA   Pellett P.E., Jenkins F.J., Ackermann M., Sarmiento M., Roizman B.;
RT   "Transcription initiation sites and nucleotide sequence of a herpes
RT   simplex virus 1 gene conserved in the Epstein-Barr virus genome and
RT   reported to affect the transport of viral glycoproteins.";
RL   J. Virol. 60:1134-1140(1986).
CC   -!- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLOCATION OF THE VIRUSES'
CC       GLYCOPROTEINS TO MEMBRANES. IT IS INVOLVED IN CAPSID MATURATION.
CC   -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES PRTP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14164; AAA45775.1; -.
DR   PIR; A26101; WMBEH8.
DR   InterPro; IPR000501; -.
DR   Pfam; PF01366; PRTP; 1.
KW   Late protein; Capsid assembly.
SQ   SEQUENCE   780 AA;  83845 MW;  AD20A1166DF4BCBE CRC64;


  Query Match             1.9%;  Score 8;  DB 1;  Length 780;
  Best Local Similarity   100.0%;  Pred. No. 2.9;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     635 LKEELARF 642


RESULT   5
PRTP_HSV11
ID   PRTP_HSV11     STANDARD;      PRT;   785 AA.
AC   P10212;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROCESSING AND TRANSPORT PROTEIN (INFECTED CELL PROTEIN 18.5) (ICP18.5
DE   PROTEIN).
GN   UL28.
OS   Herpes simplex virus (type 1 / strain 17).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
OX   NCBI_TaxID=10299;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88274327; PubMed=2839594;
RA   McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA   McNab D., Perry L.J., Scott J.E., Taylor P.;
RT   "The complete DNA sequence of the long unique region in the genome of
RT   herpes simplex virus type 1.";
RL   J. Gen. Virol. 69:1531-1574(1988).
CC   -!- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLOCATION OF THE VIRUSES'
CC       GLYCOPROTEINS TO MEMBRANES. IT IS INVOLVED IN CAPSID MATURATION.
CC   -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES PRTP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D10879; BAA01674.1; -.
DR   EMBL; X14112; CAA32321.1; -.
DR   PIR; A30085; WMBEW8.
DR   InterPro; IPR000501; -.
DR   Pfam; PF01366; PRTP; 1.
KW   Late protein; Capsid assembly.
SQ   SEQUENCE   785 AA;  85577 MW;  CFE8868BAA86F616 CRC64;


  Query Match             1.9%;  Score 8;  DB 1;  Length 785;
  Best Local Similarity   100.0%;  Pred. No. 2.9;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     640 LKEELARF 647


RESULT   6
PRTP_HSV1A
ID   PRTP_HSV1A     STANDARD;      PRT;   785 AA.
AC   P12835;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROCESSING AND TRANSPORT PROTEIN (INFECTED CELL PROTEIN 18.5) (ICP18.5
DE   PROTEIN).
GN   UL28.
OS   Herpes simplex virus (type 1 / strain Angelotti).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
OX   NCBI_TaxID=10301;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88040433; PubMed=2823222;
RA   Knopf C.W.;
RT   "The nucleotide sequence of the herpes simplex virus type 1 late gene
RT   ICP18.5 of strain Angelotti.";
RL   Nucleic Acids Res. 15:8109-8110(1987).
CC   -!- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLOCATION OF THE VIRUSES'
CC       GLYCOPROTEINS TO MEMBRANES. IT IS INVOLVED IN CAPSID MATURATION.
CC   -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES PRTP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y00453; CAA68508.1; -.
DR   PIR; A26987; WMBEK8.
DR   InterPro; IPR000501; -.
DR   Pfam; PF01366; PRTP; 1.
KW   Late protein; Capsid assembly.
SQ   SEQUENCE   785 AA;  85607 MW;  7E08C7D4B4734D1A CRC64;


  Query Match             1.9%;  Score 8;  DB 1;  Length 785;
  Best Local Similarity   100.0%;  Pred. No. 2.9;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     640 LKEELARF 647


RESULT   7
HV3H_HUMAN
ID   HV3H_HUMAN      STANDARD;      PRT;   122 AA.
AC   P01769;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   IG HEAVY CHAIN V-III REGION GA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=74175307; PubMed=4208843;
RA   Florent G., Lehman D., Putnam F.W.;
RT   "The switch point in mu heavy chains of human IgM immunoglobulins.";
RL   Biochemistry 13:2482-2498(1974).
CC   -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A WALDENSTROM'S
CC       MACROGLOBULIN.
DR   PIR; A02052; M3HUGA.
DR   HSSP; P01772; 2IG2.
DR   InterPro; IPR003006; -.
DR   Pfam; PF00047; ig; 1.
KW   Immunoglobulin V region.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   NON_TER     122    122
SQ   SEQUENCE   122 AA;  13166 MW;  74E5B6959E84100A CRC64;


  Query Match             1.6%;  Score 7;  DB 1;  Length 122;
  Best Local Similarity   100.0%;  Pred. No. 6.5;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     297 WLSVISY 303
           |||||||
Db      47 WLSVISY 53


RESULT   8
DRN1_PIG
ID   DRN1_PIG       STANDARD;      PRT;   262 AA.
AC   P11936;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   DEOXYRIBONUCLEASE I (EC 3.1.21.1) (DNASE I).
GN   DNASE1 OR DNL1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE, CARBOHYDRATE-LINKAGE SITES, AND DISULFIDE BONDS.
RC   TISSUE=Pancreas;
RX   MEDLINE=87057260; PubMed=3782104;
RA   Paudel H.K., Liao T.-H.;
RT   "Purification, characterization, and the complete amino acid sequence
RT   of porcine pancreatic deoxyribonuclease.";
RL   J. Biol. Chem. 261:16006-16011(1986).
CC   -!- FUNCTION: AMONG OTHER FUNCTIONS, SEEMS TO BE INVOLVED IN CELL
CC       DEATH BY APOPTOSIS. BINDS SPECIFICALLY TO G-ACTIN AND BLOCKS
CC       ACTIN POLYMERIZATION (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE TO 5'-
CC       PHOSPHODINUCLEOTIDE AND 5'-PHOSPHOOLIGONUCLEOTIDE END-PRODUCTS.
CC   -!- COFACTOR: DIVALENT CATIONS, PARTICULARLY CALCIUM AND MAGNESIUM.
CC   -!- SUBCELLULAR LOCATION: SECRETORY PROTEIN, STORED IN ZYMOGEN
CC       GRANULES AND FOUND IN THE NUCLEAR ENVELOPE.
CC   -!- SIMILARITY: BELONGS TO THE DNASE I FAMILY.
DR   PIR; A26324; A26324.
DR   HSSP; P00639; 1ATN.
DR   InterPro; IPR001582; -.
DR   Pfam; PF01181; DNase_I; 1.
DR   PRINTS; PR00130; DNASEI.
DR   PROSITE; PS00919; DNASE_I_1; 1.
DR   PROSITE; PS00918; DNASE_I_2; 1.
KW   Hydrolase; Endonuclease; Nuclease; Glycoprotein; Calcium;
KW   Actin-binding.
FT   DISULFID    101    104
FT   DISULFID    173    209       ESSENTIAL FOR ENZYMATIC ACTIVITY
FT                                (BY SIMILARITY).
FT   ACT_SITE     78     78       BY SIMILARITY.
FT   ACT_SITE    134    134       BY SIMILARITY.
FT   CARBOHYD     18     18       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    106    106       N-LINKED (GLCNAC. . .).
SQ   SEQUENCE   262 AA;  29447 MW;  FB63A6BDD81CD7C5 CRC64;


  Query Match             1.6%;  Score 7;  DB 1;  Length 262;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      43 DSYLYDD 49
           |||||||
Db      93 DSYLYDD 99


RESULT   9
ALF_HELPY
ID   ALF_HELPY      STANDARD;      PRT;   307 AA.
AC   P56109;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   FRUCTOSE-BISPHOSPHATE ALDOLASE (EC 4.1.2.13).
GN   FBA OR HP0176.
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=210;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=26695 / ATCC 700392;
RX   MEDLINE=97394467; PubMed=9252185;
RA   Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA   Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA   Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA   Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA   McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA   Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA   Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA   Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA   Venter J.C.;
RT   "The complete genome sequence of the gastric pathogen Helicobacter
RT   pylori.";
RL   Nature 388:539-547(1997).
CC   -!- CATALYTIC ACTIVITY: D-FRUCTOSE 1,6-BISPHOSPHATE = GLYCERONE-
CC       PHOSPHATE + D-GLYCERALDEHYDE 3-PHOSPHATE.
CC   -!- COFACTOR: ZINC (BY SIMILARITY).
CC   -!- PATHWAY: SIXTH STEP IN GLYCOLYSIS.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO CLASS II FRUCTOSE-BISPHOSPHATE ALDOLASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000538; AAD07246.1; -.
DR   TIGR; HP0176; -.
DR   InterPro; IPR000771; -.
DR   Pfam; PF01116; F_bP_aldolase; 1.
DR   PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR   PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
KW   Lyase; Glycolysis; Zinc.
FT   METAL        80     80       ZINC (BY SIMILARITY).
FT   METAL        83     83       ZINC (BY SIMILARITY).
SQ   SEQUENCE   307 AA;  33773 MW;  ADC5696C67C69A45 CRC64;


  Query Match             1.6%;  Score 7;  DB 1;  Length 307;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     402 VKGGINK 408
           |||||||
Db     245 VKGGINK 251


RESULT  10
DHB1_MOUSE
ID   DHB1_MOUSE     STANDARD;      PRT;   344 AA.
AC   P51656;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   ESTRADIOL 17 BETA-DEHYDROGENASE 1 (EC 1.1.1.62) (17-BETA-HSD 1)
DE   (17-BETA-HYDROXYSTEROID DEHYDROGENASE 1).
GN   HSD17B1 OR EDH17B1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=96195656; PubMed=8612620;
RA   Nokelainen P., Puranen T., Peltoketo H., Orava M., Vihko P.,
RA   Vihko R.;
RT   "Molecular cloning of mouse 17 beta-hydroxysteroid dehydrogenase type
RT   1 and characterization of enzyme activity.";
RL   Eur. J. Biochem. 236:482-490(1996).
CC   -!- FUNCTION: FAVORS THE REDUCTION OF ESTROGENS AND ANDROGENS.
CC       USES PREFERENTIALLY NADH.
CC   -!- CATALYTIC ACTIVITY: ESTRADIOL-17-BETA + NAD(+) = ESTRONE +
CC       NADH.
CC   -!- PATHWAY: BIOSYNTHESIS OF ESTROGENS.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC       (SDR) FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X89627; CAA61770.1; -.
DR   HSSP; P14061; 1FDW.
DR   MGD; MGI:105077; Hsd17b1.
DR   InterPro; IPR002198; -.
DR   Pfam; PF00106; adh_short; 1.
DR   PROSITE; PS00061; ADH_SHORT; 1.
KW   Steroid biosynthesis; Oxidoreductase; NAD; Multigene family.
FT   NP_BIND       3     32       NAD (BY SIMILARITY).
FT   ACT_SITE    143    143       SUBSTRATE BINDING (BY SIMILARITY).
FT   ACT_SITE    156    156       BY SIMILARITY.
SQ   SEQUENCE   344 AA;  36785 MW;  0335C473F453C14A CRC64;


  Query Match             1.6%;  Score 7;  DB 1;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     220 PGSLEIL 226
           |||||||
Db      57 PGSLEIL 63


RESULT  11
DHB1_RAT
ID   DHB1_RAT       STANDARD;      PRT;   344 AA.
AC   P51657;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   ESTRADIOL 17 BETA-DEHYDROGENASE 1 (EC 1.1.1.62) (17-BETA-HSD 1)
DE   (17-BETA-HYDROXYSTEROID DEHYDROGENASE 1).
GN   HSD17B1 OR EDH17B1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Ovary;
RX   MEDLINE=95009707; PubMed=7925110;
RA   Ghersevich S., Nokelainen P., Poutanen M., Orava M.,
RA   Autio-Harmainen H., Rajaniemi H., Vihko R.;
RT   "Rat 17 beta-hydroxysteroid dehydrogenase type 1: primary structure
RT   and regulation of enzyme expression in rat ovary by
RT   diethylstilbestrol and gonadotropins in vivo.";
RL   Endocrinology 135:1477-1487(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney;
RX   MEDLINE=96198748; PubMed=8612487;
RA   Akinola L.A., Poutanen M., Vihko R.;
RT   "Cloning of rat 17 beta-hydroxysteroid dehydrogenase type 2 and
```

```
RT   characterization of tissue distribution and catalytic activity of rat
RT   type 1 and type 2 enzymes.";
RL   Endocrinology 137:1572-1579(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Testis;
RX   MEDLINE=98201618; PubMed=9524272;
RA   Akinola L.A., Poutanen M., Peltoketo H., Vihko R., Vihko P.;
RT   "Characterization of rat 17 beta-hydroxysteroid dehydrogenase type 1
RT   gene and mRNA transcripts.";
RL   Gene 208:229-238(1998).
CC   -!- FUNCTION: FAVORS THE REDUCTION OF ESTROGENS AND ANDROGENS.
CC       USES PREFERENTIALLY NADH.
CC   -!- CATALYTIC ACTIVITY: ESTRADIOL-17-BETA + NAD(+) = ESTRONE +
CC       NADH.
CC   -!- PATHWAY: BIOSYNTHESIS OF ESTROGENS.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC       (SDR) FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X78811; CAA55389.1; -.
DR   EMBL; X97754; CAA66349.1; -.
DR   EMBL; X98038; CAA66657.1; -.
DR   HSSP; P14061; 1FDW.
DR   InterPro; IPR002198; -.
DR   Pfam; PF00106; adh_short; 1.
DR   PROSITE; PS00061; ADH_SHORT; 1.
KW   Steroid biosynthesis; Oxidoreductase; NAD; Multigene family.
FT   NP_BIND      3     32       NAD (BY SIMILARITY).
FT   ACT_SITE   143    143       SUBSTRATE BINDING (BY SIMILARITY).
FT   ACT_SITE   156    156       BY SIMILARITY.
SQ   SEQUENCE   344 AA;  36967 MW;  CB5FC139FBB3995E CRC64;


  Query Match             1.6%;  Score 7;  DB 1;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    220 PGSLEIL 226
          |||||||
Db     57 PGSLEIL 63


RESULT  12
PERM_ECOLI
ID   PERM_ECOLI     STANDARD;      PRT;   353 AA.
AC   P77406; P71230;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PUTATIVE PERMEASE PERM.
GN   PERM.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RA   Aiba H., Baba T., Fujita K., Hayashi K., Honjo A., Horiuchi T.,
RA   Ikemoto K., Inada T., Isono K., Isono S., Itoh T., Kanai K., Kasai H.,
RA   Kashimoto K., Kim S., Kimura S., Kitagawa M., Kitakawa M., Makino K.,
RA   Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y.,
RA   Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA   Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 279-353 FROM N.A.
RC   STRAIN=K12;
RA   Andrews S.C., McClay J., Ambler A., Quail M., Berks B.C., Guest J.R.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE UPF0118 (PERM) FAMILY. STRONG, TO
CC       H.INFLUENZAE HI0237/38.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000335; AAC75546.1; -.
DR   EMBL; D90878; BAA16382.1; -.
DR   EMBL; M63654; AAB88575.1; ALT_SEQ.
DR   EcoGene; EG14221; perM.
DR   InterPro; IPR002549; -.
DR   Pfam; PF01594; DUF20; 1.
KW   Transmembrane; Transport.
FT   TRANSMEM     19     39       POTENTIAL.
FT   TRANSMEM     72     92       POTENTIAL.
FT   TRANSMEM    156    176       POTENTIAL.
FT   TRANSMEM    217    237       POTENTIAL.
FT   TRANSMEM    240    260       POTENTIAL.
FT   TRANSMEM    281    301       POTENTIAL.
FT   TRANSMEM    310    330       POTENTIAL.
SQ   SEQUENCE   353 AA;  39194 MW;  E1CB8EB1E242FC3B CRC64;


  Query Match             1.6%;  Score 7;  DB 1;  Length 353;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    235 LTIAVYL 241
          |||||||
Db    160 LTIAVYL 166


RESULT  13
SUC2_ARCFU
ID   SUC2_ARCFU     STANDARD;      PRT;   380 AA.
AC   O28097;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SUCCINYL-COA SYNTHETASE BETA CHAIN 2 (EC 6.2.1.5) (SCS-BETA 2).
GN   SUCC2 OR AF2186.
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC   Archaeoglobus.
OX   NCBI_TaxID=2234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX   MEDLINE=98049343; PubMed=9389475;
RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
```

```
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
RT   reducing archaeon Archaeoglobus fulgidus.";
RL   Nature 390:364-370(1997).
CC   -!- CATALYTIC ACTIVITY: SUCCINATE + COA + ATP = SUCCINYL-COA + ADP +
CC       ORTHOPHOSPHATE.
CC   -!- SIMILARITY: TO OTHER BETA SUBUNITS OF SUCCINYL-COA SYNTHETASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000954; AAB89066.1; -.
DR   TIGR; AF2186; -.
DR   InterPro; IPR000303; -.
DR   Pfam; PF00549; ligase-CoA; 1.
DR   PROSITE; PS01217; SUCCINYL_COA_LIG_3; 1.
KW   Ligase; Tricarboxylic acid cycle.
SQ   SEQUENCE   380 AA;  41998 MW;  538FDE4CE286DB96 CRC64;


  Query Match             1.6%;  Score 7;  DB 1;  Length 380;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     410 TNEEEGR 416
           |||||||
Db     341 TNEEEGR 347


RESULT  14
K1CS_BOVIN
ID   K1CS_BOVIN     STANDARD;      PRT;   399 AA.
AC   P08728;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   KERATIN, TYPE I CYTOSKELETAL 19 (CYTOKERATIN 19) (K19) (CK 19).
GN   KRT19.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87004553; PubMed=2428612;
RA   Bader B.L., Magin T.M., Hatzfeld M., Franke W.W.;
RT   "Amino acid sequence and gene organization of cytokeratin no. 19, an
RT   exceptional tail-less intermediate filament protein.";
RL   EMBO J. 5:1865-1875(1986).
CC   -!- SUBUNIT: HETEROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC   -!- DOMAIN: THIS KERATIN DIFFERS FROM ALL OTHER IF PROTEINS IN LACKING
CC       THE C-TERMINAL TAIL DOMAIN.
CC   -!- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
CC       MICROFIBRILLAR KERATIN, I (ACIDIC) AND II (NEUTRAL TO BASIC)
CC       (40-55 AND 56-70 KILODALTONS, RESPECTIVELY).
CC   -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04152; CAA27770.1; -.
DR   EMBL; X04198; CAA27770.1; JOINED.
DR   PIR; A25470; A25470.
DR   InterPro; IPR001664; -.
DR   InterPro; IPR002957; -.
DR   Pfam; PF00038; filament; 1.
DR   PRINTS; PR01248; TYPE1KERATIN.
DR   PROSITE; PS00226; IF; 1.
KW   Intermediate filament; Coiled coil; Heptad repeat pattern; Keratin.
FT   DOMAIN        1     78       HEAD.
FT   DOMAIN       79    386       ROD.
FT   DOMAIN      387    399       ROD-LIKE HELICAL TAIL.
FT   DOMAIN       79    114       COIL 1A.
FT   DOMAIN      115    132       LINKER 1.
FT   DOMAIN      133    224       COIL 1B.
FT   DOMAIN      225    247       LINKER 12.
FT   DOMAIN      248    386       COIL 2.
FT   SITE        266    266       STUTTER.
FT   SITE        326    326       STUTTER.
SQ   SEQUENCE   399 AA;  43885 MW;  8F2E12149A7B99F2 CRC64;


  Query Match             1.6%;  Score 7;  DB 1;  Length 399;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     106 GLKEELA 112
           |||||||
Db     205 GLKEELA 211


RESULT  15
K1CS_HUMAN
ID   K1CS_HUMAN     STANDARD;      PRT;   400 AA.
AC   P08727;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   KERATIN, TYPE I CYTOSKELETAL 19 (CYTOKERATIN 19) (K19) (CK 19).
GN   KRT19.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89235250; PubMed=2469734;
RA   Stasiak P.C., Purkis P.E., Leigh I.M., Lane E.B.;
RT   "Keratin 19: predicted amino acid sequence and broad tissue
RT   distribution suggest it evolved from keratinocyte keratins.";
RL   J. Invest. Dermatol. 92:707-716(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=88096504; PubMed=2447559;
RA   Stasiak P.C., Lane E.B.;
RT   "Sequence of cDNA coding for human keratin 19.";
RL   Nucleic Acids Res. 15:10058-10058(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88124986; PubMed=2448790;
RA   Eckert R.L.;
RT   "Sequence of the human 40-kDa keratin reveals an unusual structure
RT   with very high sequence identity to the corresponding bovine
RT   keratin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 85:1114-1118(1988).
```

```
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20090660; PubMed=10623642;
RA   Whittock N.V., Eady R.A., McGrath J.A.;
RT   "Genomic organization and amplification of the human keratin 15 and 19
RT   genes.";
RL   Biochem. Biophys. Res. Commun. 267:462-465(2000).
RN   [5]
RP   SEQUENCE OF 25-31; 151-158 AND 227-237.
RC   TISSUE=Keratinocytes;
RX   MEDLINE=93162043; PubMed=1286667;
RA   Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
RA   Vandekerckhove J.;
RT   "Microsequences of 145 proteins recorded in the two-dimensional gel
RT   protein database of normal human epidermal keratinocytes.";
RL   Electrophoresis 13:960-969(1992).
CC   -!- FUNCTION: THERE ARE TWO TYPES OF CYTOSKELETAL AND MICROFIBRILLAR
CC       KERATIN: I (ACIDIC; 40-55 KDA) [K9 TO K20] AND II (NEUTRAL TO
CC       BASIC; 56-70 KDA) [K1 TO K8]. BOTH A BASIC AND AN ACIDIC KERATIN
CC       ARE REQUIRED FOR FILAMENT ASSEMBLY.
CC   -!- SUBUNIT: HETEROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC   -!- DOMAIN: THIS KERATIN DIFFERS FROM ALL OTHER IF PROTEINS IN LACKING
CC       THE C-TERMINAL TAIL DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y00503; CAA68556.1; -.
DR   EMBL; J03607; AAA36044.1; -.
DR   EMBL; AF202321; AAF27048.1; -.
DR   PIR; A60779; KRHU9.
DR   SWISS-2DPAGE; P08727; HUMAN.
DR   Aarhus/Ghent-2DPAGE; 8216; IEF.
DR   MIM; 148020; -.
DR   InterPro; IPR001664; -.
DR   InterPro; IPR002957; -.
DR   Pfam; PF00038; filament; 1.
DR   PRINTS; PR01248; TYPE1KERATIN.
DR   PROSITE; PS00226; IF; 1.
KW   Intermediate filament; Coiled coil; Heptad repeat pattern; Keratin.
FT   DOMAIN        1     79       HEAD.
FT   DOMAIN       80    387       ROD.
FT   DOMAIN      388    400       ROD-LIKE HELICAL TAIL.
FT   DOMAIN       80    115       COIL 1A.
FT   DOMAIN      116    133       LINKER 1.
FT   DOMAIN      134    225       COIL 1B.
FT   DOMAIN      226    248       LINKER 12.
FT   DOMAIN      249    387       COIL 2.
FT   SITE        267    267       STUTTER.
FT   SITE        327    327       STUTTER.
FT   CONFLICT     76     77       LA -> H (IN REF. 2).
FT   CONFLICT    342    342       I -> Y (IN REF. 2).
FT   CONFLICT    350    350       A -> G (IN REF. 3).
SQ   SEQUENCE   400 AA;  44106 MW;  1EBF9AC82F09CAB5 CRC64;


  Query Match             1.6%;  Score 7;  DB 1;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    106 GLKEELA 112
          |||||||
Db    206 GLKEELA 212


Search completed: October 30, 2001, 15:33:51

Job time: 217 sec
```

THIS PAGE BLANK (USPTO)

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 15:29:54 ; Search time 22.67 Seconds
(without alignments)
2486.191 Million cell updates/sec

Title: US-09-619-049-855
Perfect score: 426
Sequence: 1 MRLPWVIFCTVLLLIFTNNS.....NKFTNEEEGRINCINFDSQT 426

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SPTREMBL_16:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	426	100.0	426	5	Q9VSE7	Q9vse7 drosophila
2	8	1.9	179	14	Q69093	Q69093 herpes simp
3	8	1.9	266	14	Q04459	Q04459 herpesvirus
4	8	1.9	270	3	O94006	O94006 candida alb
5	8	1.9	362	14	Q69261	Q69261 equine herp
6	8	1.9	482	5	Q9VRN2	Q9vrn2 drosophila
7	8	1.9	513	5	Q9W0R4	Q9w0r4 drosophila
8	8	1.9	514	5	O97148	O97148 drosophila
9	8	1.9	517	5	Q9GT56	Q9gt56 drosophila
10	8	1.9	517	5	Q9GT55	Q9gt55 drosophila
11	8	1.9	517	5	Q9GT53	Q9gt53 drosophila
12	8	1.9	517	5	Q9GT52	Q9gt52 drosophila
13	8	1.9	517	5	Q9GT51	Q9gt51 drosophila
14	8	1.9	773	14	O39274	O39274 equine herp
15	8	1.9	785	14	P89451	P89451 herpes simp
16	8	1.9	3738	10	O65645	O65645 arabidopsis
17	8	1.9	3795	10	Q9SL52	Q9sl52 arabidopsis
18	7	1.6	76	2	Q9K718	Q9k718 bacillus ha
19	7	1.6	104	4	Q9P1Y3	Q9ply3 homo sapien
20	7	1.6	116	2	Q9F9T8	Q9f9t8 edta-degrad
21	7	1.6	118	11	O08926	O08926 mus musculu
22	7	1.6	126	10	Q42078	Q42078 arabidopsis
23	7	1.6	137	2	Q9PCT9	Q9pct9 xylella fas
24	7	1.6	142	5	O16896	O16896 caenorhabdi
25	7	1.6	145	2	Q9L8C2	Q918c2 sorangium c
26	7	1.6	186	8	Q34547	Q34547 glyphinaphi
27	7	1.6	193	2	Q9K526	Q9k526 mycobacteri
28	7	1.6	198	8	Q9MEU5	Q9meu5 schistosoma
29	7	1.6	208	4	Q9P1Y4	Q9ply4 homo sapien
30	7	1.6	214	2	O68758	O68758 yersinia pe
31	7	1.6	244	4	Q16195	Q16195 homo sapien
32	7	1.6	260	10	Q9FQ25	Q9fq25 arabidopsis
33	7	1.6	308	5	Q9VLH0	Q9vlh0 drosophila
34	7	1.6	311	10	O64836	O64836 arabidopsis
35	7	1.6	328	2	O06298	O06298 mycobacteri
36	7	1.6	329	2	O68555	O68555 corynebacte
37	7	1.6	329	2	O69274	O69274 corynebacte
38	7	1.6	336	5	Q9U0G9	Q9u0g9 plasmodium
39	7	1.6	341	2	O30880	O30880 pseudomonas
40	7	1.6	342	5	Q9VD17	Q9vd17 drosophila
41	7	1.6	351	14	Q9E9F4	Q9e9f4 human t-cel
42	7	1.6	353	2	Q9Z8K9	Q9z8k9 chlamydia p
43	7	1.6	353	2	Q9K270	Q9k270 chlamydia p
44	7	1.6	367	5	O17973	O17973 caenorhabdi
45	7	1.6	372	2	O05804	O05804 mycobacteri
46	7	1.6	372	10	Q9LHK0	Q9lhk0 arabidopsis
47	7	1.6	379	2	O87181	O87181 streptococc
48	7	1.6	380	1	O29694	O29694 archaeoglob
49	7	1.6	385	4	O76067	O76067 homo sapien
50	7	1.6	396	11	O88634	O88634 mus musculu
51	7	1.6	398	1	Q9P9M8	Q9p9m8 pyrococcus
52	7	1.6	401	1	O59044	O59044 pyrococcus
53	7	1.6	458	10	Q9LVF3	Q9lvf3 arabidopsis
54	7	1.6	471	11	Q9R1F9	Q9r1f9 mus musculu
55	7	1.6	473	4	Q16402	Q16402 homo sapien
56	7	1.6	473	4	Q9UBG8	Q9ubg8 homo sapien
57	7	1.6	480	10	O23486	O23486 arabidopsis
58	7	1.6	492	5	Q9W0V7	Q9w0v7 drosophila
59	7	1.6	497	8	Q9MLH5	Q9mlh5 paragonimus
60	7	1.6	517	5	Q9GT54	Q9gt54 drosophila
61	7	1.6	517	5	Q9GT50	Q9gt50 drosophila
62	7	1.6	521	2	Q9RYN3	Q9ryn3 deinococcus
63	7	1.6	554	3	Q9US61	Q9us61 schizosacch
64	7	1.6	554	3	Q9US60	Q9us60 schizosacch
65	7	1.6	588	4	Q9NZH1	Q9nzh1 homo sapien
66	7	1.6	600	3	Q9HGI2	Q9hgi2 neurospora
67	7	1.6	709	4	O75619	O75619 homo sapien
68	7	1.6	712	4	O15060	O15060 homo sapien
69	7	1.6	723	5	Q9U4H6	Q9u4h6 caenorhabdi
70	7	1.6	738	2	Q9X3V8	Q9x3v8 zymomonas m
71	7	1.6	743	5	O17305	O17305 caenorhabdi
72	7	1.6	769	5	Q9N3Y9	Q9n3y9 caenorhabdi
73	7	1.6	791	2	O83389	O83389 treponema p
74	7	1.6	792	5	O45918	O45918 caenorhabdi
75	7	1.6	943	4	O94960	O94960 homo sapien
76	7	1.6	1042	2	Q9KXM6	Q9kxm6 streptomyce
77	7	1.6	1336	2	O52800	O52800 amycolatops
78	7	1.6	1467	10	Q9M033	Q9m033 arabidopsis
79	7	1.6	1960	5	Q9U0K8	Q9u0k8 plasmodium
80	7	1.6	2010	5	O96171	O96171 plasmodium
81	7	1.6	2241	5	Q19433	Q19433 caenorhabdi
82	7	1.6	3085	14	O70710	O70710 infectious
83	7	1.6	3225	2	Q9PKM6	Q9pkm6 chlamydia m
84	6	1.4	30	10	Q9SJL3	Q9sjl3 arabidopsis
85	6	1.4	34	6	O46539	O46539 sus scrofa
86	6	1.4	44	5	Q9TWW4	Q9tww4 lymnaea sta
87	6	1.4	47	2	O50722	O50722 borrelia bu
88	6	1.4	47	11	Q61346	Q61346 mus musculu
89	6	1.4	49	6	O97717	O97717 bos taurus
90	6	1.4	52	9	Q9T1B6	Q9t1b6 bacteriopha
91	6	1.4	57	5	Q9NLI8	Q9nli8 leishmania
92	6	1.4	69	5	O61759	O61759 caenorhabdi

93	6	1.4	76	5	Q9XTX9	Q9xtx9 caenorhabdi
94	6	1.4	78	7	Q9MWR7	Q9mwr7 leopardus w
95	6	1.4	78	14	Q9YRA3	Q9yra3 porcine tra
96	6	1.4	79	7	O19378	O19378 felis silve
97	6	1.4	79	7	O19379	O19379 felis silve
98	6	1.4	79	7	O19380	O19380 felis silve
99	6	1.4	79	7	O19381	O19381 felis silve
100	6	1.4	79	7	O19382	O19382 felis silve
101	6	1.4	79	7	O19383	O19383 felis silve
102	6	1.4	79	7	O19384	O19384 felis silve
103	6	1.4	79	7	O19385	O19385 felis silve
104	6	1.4	79	7	O19386	O19386 felis silve
105	6	1.4	79	7	O19387	O19387 felis silve
106	6	1.4	79	7	O19388	O19388 felis silve
107	6	1.4	79	7	O19389	O19389 felis silve
108	6	1.4	79	7	O19390	O19390 felis silve
109	6	1.4	79	7	O19391	O19391 felis silve
110	6	1.4	79	7	O19392	O19392 felis silve
111	6	1.4	79	7	O19393	O19393 felis silve
112	6	1.4	79	7	O19394	O19394 felis silve
113	6	1.4	79	7	O19395	O19395 felis silve
114	6	1.4	79	7	O19428	O19428 oncifelis g
115	6	1.4	79	7	O19435	O19435 felis silve
116	6	1.4	79	7	Q9MX20	Q9mx20 leopardus p
117	6	1.4	79	7	Q9MX18	Q9mx18 leopardus p
118	6	1.4	79	7	Q9MX17	Q9mx17 leopardus p
119	6	1.4	79	7	Q9MX16	Q9mx16 leopardus p
120	6	1.4	79	7	Q9MX15	Q9mx15 leopardus p
121	6	1.4	79	7	Q9MX14	Q9mx14 leopardus p
122	6	1.4	79	7	Q9MX13	Q9mx13 leopardus p
123	6	1.4	79	7	Q9MX12	Q9mx12 leopardus p
124	6	1.4	79	7	Q9MX11	Q9mx11 leopardus p
125	6	1.4	79	7	Q9MX10	Q9mx10 leopardus p
126	6	1.4	79	7	Q9MX09	Q9mx09 leopardus p
127	6	1.4	79	7	Q9MX08	Q9mx08 leopardus p
128	6	1.4	79	7	Q9MX07	Q9mx07 leopardus p
129	6	1.4	79	7	Q9MX06	Q9mx06 leopardus p
130	6	1.4	79	7	Q9MX05	Q9mx05 leopardus p
131	6	1.4	79	7	Q9MX04	Q9mx04 leopardus p
132	6	1.4	79	7	Q9MX03	Q9mx03 leopardus p
133	6	1.4	79	7	Q9MX02	Q9mx02 leopardus p
134	6	1.4	79	7	Q9MX01	Q9mx01 leopardus p
135	6	1.4	79	7	Q9MX00	Q9mx00 leopardus p
136	6	1.4	79	7	Q9MWZ9	Q9mwz9 leopardus p
137	6	1.4	79	7	Q9MWZ5	Q9mwz5 leopardus p
138	6	1.4	79	7	Q9MWZ4	Q9mwz4 leopardus p
139	6	1.4	79	7	Q9MWZ3	Q9mwz3 leopardus p
140	6	1.4	79	7	Q9MWZ2	Q9mwz2 leopardus p
141	6	1.4	79	7	Q9MWZ1	Q9mwz1 leopardus p
142	6	1.4	79	7	Q9MWZ0	Q9mwz0 leopardus p
143	6	1.4	79	7	Q9MWY9	Q9mwy9 leopardus p
144	6	1.4	79	7	Q9MWY8	Q9mwy8 leopardus p
145	6	1.4	79	7	Q9MWY7	Q9mwy7 leopardus p
146	6	1.4	79	7	Q9MWY6	Q9mwy6 leopardus p
147	6	1.4	79	7	Q9MWY5	Q9mwy5 leopardus p
148	6	1.4	79	7	Q9MWY4	Q9mwy4 leopardus p
149	6	1.4	79	7	Q9MWY3	Q9mwy3 leopardus p
150	6	1.4	79	7	Q9MWY2	Q9mwy2 leopardus p
151	6	1.4	79	7	Q9MWY1	Q9mwy1 leopardus p
152	6	1.4	79	7	Q9MWY0	Q9mwy0 leopardus p
153	6	1.4	79	7	Q9MWX8	Q9mwx8 leopardus p
154	6	1.4	79	7	Q9MWX6	Q9mwx6 leopardus p
155	6	1.4	79	7	Q9MWX5	Q9mwx5 leopardus p
156	6	1.4	79	7	Q9MWX4	Q9mwx4 leopardus p
157	6	1.4	79	7	Q9MWX3	Q9mwx3 leopardus p
158	6	1.4	79	7	Q9MWX2	Q9mwx2 leopardus p
159	6	1.4	79	7	Q9MWX1	Q9mwx1 leopardus p
160	6	1.4	79	7	Q9MWX0	Q9mwx0 leopardus p
161	6	1.4	79	7	Q9MWW9	Q9mww9 leopardus p
162	6	1.4	79	7	Q9MWW8	Q9mww8 leopardus p
163	6	1.4	79	7	Q9MWW7	Q9mww7 leopardus p
164	6	1.4	79	7	Q9MWT0	Q9mwt0 leopardus p
165	6	1.4	79	7	Q9MWS9	Q9mws9 leopardus p
166	6	1.4	79	7	Q9MWS3	Q9mws3 leopardus p
167	6	1.4	79	7	Q9MWS2	Q9mws2 leopardus p
168	6	1.4	79	7	Q9MWS1	Q9mws1 leopardus p
169	6	1.4	79	7	Q9MWS0	Q9mws0 leopardus p
170	6	1.4	79	7	Q9MWR9	Q9mwr9 leopardus p
171	6	1.4	79	7	Q9MWR8	Q9mwr8 leopardus p
172	6	1.4	79	7	Q9MWR6	Q9mwr6 leopardus w
173	6	1.4	79	7	Q9MWR5	Q9mwr5 leopardus w
174	6	1.4	79	7	Q9MWR4	Q9mwr4 leopardus w
175	6	1.4	79	7	Q9MWR3	Q9mwr3 leopardus w
176	6	1.4	79	7	Q9MWR2	Q9mwr2 leopardus w
177	6	1.4	79	7	Q9MWR1	Q9mwr1 leopardus w
178	6	1.4	79	7	Q9MWQ9	Q9mwq9 leopardus w
179	6	1.4	79	7	Q9MWQ8	Q9mwq8 leopardus w
180	6	1.4	79	7	Q9MWQ7	Q9mwq7 leopardus w
181	6	1.4	79	7	Q9MWQ6	Q9mwq6 leopardus w
182	6	1.4	79	7	Q9MWQ5	Q9mwq5 leopardus w
183	6	1.4	79	7	Q9MWQ4	Q9mwq4 leopardus w
184	6	1.4	79	7	Q9MWQ3	Q9mwq3 leopardus w
185	6	1.4	79	7	Q9MWQ2	Q9mwq2 leopardus w
186	6	1.4	79	7	Q9MWQ1	Q9mwq1 leopardus w
187	6	1.4	79	7	Q9MWQ0	Q9mwq0 leopardus w
188	6	1.4	79	7	Q9MWP8	Q9mwp8 leopardus w
189	6	1.4	79	7	Q9MWP7	Q9mwp7 leopardus w
190	6	1.4	79	7	Q9MWP6	Q9mwp6 leopardus w
191	6	1.4	79	7	Q9MWP5	Q9mwp5 leopardus w
192	6	1.4	79	7	Q9MWP4	Q9mwp4 leopardus w
193	6	1.4	79	7	Q9MWP3	Q9mwp3 leopardus w
194	6	1.4	79	7	Q9MWP2	Q9mwp2 leopardus w
195	6	1.4	79	7	Q9MWP1	Q9mwp1 leopardus w
196	6	1.4	79	7	Q9MWP0	Q9mwp0 leopardus w
197	6	1.4	79	7	Q9MWN9	Q9mwn9 leopardus w
198	6	1.4	80	1	O26182	O26182 methanobact
199	6	1.4	84	10	O48816	O48816 arabidopsis
200	6	1.4	86	13	Q9W717	Q9w717 naja naja (
201	6	1.4	87	10	Q9STV2	Q9stv2 arabidopsis
202	6	1.4	88	10	Q9LFZ0	Q9lfz0 arabidopsis
203	6	1.4	97	2	Q9EU76	Q9eu76 borrelia lo
204	6	1.4	101	4	Q14476	Q14476 homo sapien
205	6	1.4	103	6	Q9N110	Q9n110 ovis aries
206	6	1.4	104	2	Q9X5Y8	Q9x5y8 rickettsia
207	6	1.4	105	5	Q9GWC1	Q9gwc1 leishmania
208	6	1.4	106	14	O41137	O41137 paramecium
209	6	1.4	106	14	O41166	O41166 paramecium
210	6	1.4	107	8	Q9XMB6	Q9xmb6 bombus sp.
211	6	1.4	108	2	Q9X311	Q9x311 bacillus an
212	6	1.4	108	8	Q9ZYN1	Q9zyn1 tockus eryt
213	6	1.4	108	11	Q9QZ94	Q9qz94 cavia porce
214	6	1.4	108	14	O56115	O56115 human immun
215	6	1.4	109	2	Q60230	Q60230 escherichia
216	6	1.4	110	4	O75920	O75920 homo sapien
217	6	1.4	111	2	P95634	P95634 rickettsia
218	6	1.4	111	8	Q9TDI6	Q9tdi6 bombus tern
219	6	1.4	113	2	Q54836	Q54836 streptococc
220	6	1.4	113	2	Q9WYJ6	Q9wyj6 thermotoga
221	6	1.4	115	1	Q9YAF6	Q9yaf6 aeropyrum p
222	6	1.4	115	2	O84235	O84235 chlamydia t
223	6	1.4	116	1	Q9YDK2	Q9ydk2 aeropyrum p
224	6	1.4	118	2	Q9KP05	Q9kp05 vibrio chol
225	6	1.4	118	5	Q9VI56	Q9vi56 drosophila
226	6	1.4	119	11	Q9WVE7	Q9wve7 rattus norv
227	6	1.4	120	2	Q9RQM9	Q9rqm9 vibrio chol
228	6	1.4	122	14	Q98539	Q98539 paramecium
229	6	1.4	122	14	Q9YXN7	Q9yxn7 human immun
230	6	1.4	123	2	Q9PKG2	Q9pkg2 chlamydia m
231	6	1.4	123	2	Q9KVJ1	Q9kvj1 vibrio chol
232	6	1.4	123	4	Q14474	Q14474 homo sapien
233	6	1.4	125	6	Q28221	Q28221 cebus albif
234	6	1.4	129	4	Q9UNL6	Q9unl6 homo sapien
235	6	1.4	129	10	Q9SW07	Q9sw07 arabidopsis
236	6	1.4	133	10	Q9SUX3	Q9sux3 arabidopsis
237	6	1.4	135	11	Q63066	Q63066 rattus norv
238	6	1.4	137	4	Q9HBN2	Q9hbn2 homo sapien

239	6	1.4	138	4	Q9H424	Q9h424 homo sapien
240	6	1.4	138	6	Q9GLX3	Q9glx3 chiropotes
241	6	1.4	140	4	Q9HBK2	Q9hbk2 homo sapien
242	6	1.4	140	11	Q61600	Q61600 mus musculu
243	6	1.4	141	2	O68461	O68461 rickettsia
244	6	1.4	141	2	O68462	O68462 rickettsia
245	6	1.4	141	2	O68463	O68463 rickettsia
246	6	1.4	141	13	Q9PV28	Q9pv28 osteoglossu
247	6	1.4	142	2	Q9RHB5	Q9rhb5 rickettsia
248	6	1.4	144	10	Q9ST76	Q9st76 oryza sativ
249	6	1.4	145	2	Q52668	Q52668 rickettsia
250	6	1.4	145	6	Q9GJY8	Q9gjy8 callicebus
251	6	1.4	146	10	Q9FLU3	Q9flu3 arabidopsis
252	6	1.4	147	6	Q28779	Q28779 pan paniscu
253	6	1.4	147	6	Q27940	Q27940 aotus azara
254	6	1.4	147	6	Q03903	Q03903 macaca mula
255	6	1.4	147	6	Q9GLX7	Q9glx7 saimiri sci
256	6	1.4	147	6	Q9GLX6	Q9glx6 saimiri ust
257	6	1.4	147	6	Q9GLX5	Q9glx5 aotus nancy
258	6	1.4	147	6	Q9GLX4	Q9glx4 callicebus
259	6	1.4	147	6	Q9GJS7	Q9gjs7 callithrix
260	6	1.4	147	11	O88754	O88754 rattus norv
261	6	1.4	148	10	Q9LDQ9	Q9ldq9 chara coral
262	6	1.4	149	10	O82018	O82018 mougeotia s
263	6	1.4	151	5	Q9XXB9	Q9xxb9 caenorhabdi
264	6	1.4	151	13	Q9I8E5	Q9i8e5 fugu rubrip
265	6	1.4	152	4	Q14491	Q14491 homo sapien
266	6	1.4	155	4	Q14403	Q14403 homo sapien
267	6	1.4	155	8	O99320	O99320 cyparium co
268	6	1.4	156	3	Q9P640	Q9p640 neurospora
269	6	1.4	160	13	Q9W6W1	Q9w6w1 latimeria c
270	6	1.4	163	8	Q02662	Q02662 podospora a
271	6	1.4	168	2	Q9Z9V1	Q9z9v1 bacillus ha
272	6	1.4	168	5	Q22933	Q22933 caenorhabdi
273	6	1.4	169	2	O51419	O51419 borrelia bu
274	6	1.4	169	2	Q9WZE5	Q9wze5 thermotoga
275	6	1.4	170	2	O06078	O06078 mycobacteri
276	6	1.4	170	2	Q9Z956	Q9z956 chlamydia p
277	6	1.4	170	2	Q9JSI8	Q9jsi8 chlamydia p
278	6	1.4	171	2	O53704	O53704 mycobacteri
279	6	1.4	171	2	Q9RI95	Q9ri95 streptomyce
280	6	1.4	171	2	Q9L0G8	Q9l0g8 streptomyce
281	6	1.4	171	4	Q9NUW7	Q9nuw7 homo sapien
282	6	1.4	172	1	O58034	O58034 pyrococcus
283	6	1.4	172	5	Q9VS77	Q9vs77 drosophila
284	6	1.4	172	14	Q9PXK8	Q9pxk8 feline immu
285	6	1.4	175	3	Q03323	Q03323 saccharomyc
286	6	1.4	175	10	Q9LTY8	Q9lty8 arabidopsis
287	6	1.4	175	14	P88229	P88229 human immun
288	6	1.4	175	14	P88226	P88226 human immun
289	6	1.4	176	10	Q9LZ14	Q9lz14 arabidopsis
290	6	1.4	177	1	Q9HLH1	Q9hlh1 thermoplasm
291	6	1.4	177	2	P95630	P95630 rickettsia
292	6	1.4	177	2	Q52768	Q52768 rickettsia
293	6	1.4	177	2	Q9EYV6	Q9eyv6 rhizobium f
294	6	1.4	178	2	P95640	P95640 rickettsia
295	6	1.4	178	8	O99361	O99361 tetilla sp.
296	6	1.4	179	14	Q86664	Q86664 feline herp
297	6	1.4	180	11	Q9WUU5	Q9wuu5 mus musculu
298	6	1.4	180	14	Q9DHN7	Q9dhn7 yaba-like d
299	6	1.4	186	1	P95969	P95969 sulfolobus
300	6	1.4	186	5	Q19670	Q19670 caenorhabdi
301	6	1.4	187	10	Q40642	Q40642 oryza sativ
302	6	1.4	188	1	Q9HMZ2	Q9hmz2 halobacteri
303	6	1.4	188	2	Q9ZN62	Q9zn62 prevotella
304	6	1.4	188	4	Q9NZ38	Q9nz38 homo sapien
305	6	1.4	188	8	Q96177	Q96177 eulimnogamm
306	6	1.4	189	2	Q52670	Q52670 rickettsia
307	6	1.4	192	2	Q9F0Q0	Q9f0q0 rickettsia
308	6	1.4	192	5	Q9V818	Q9v818 drosophila
309	6	1.4	192	14	Q9DVS2	Q9dvs2 plutella xy
310	6	1.4	193	14	Q9QJ59	Q9qj59 human herpe
311	6	1.4	194	2	Q54912	Q54912 streptococc
312	6	1.4	194	5	Q9NAM3	Q9nam3 caenorhabdi
313	6	1.4	195	3	O94561	O94561 schizosacch
314	6	1.4	195	8	Q9G727	Q9g727 eulimnogamm
315	6	1.4	195	11	O55237	O55237 mus musculu
316	6	1.4	196	2	Q52896	Q52896 rickettsia
317	6	1.4	196	2	Q9LAR2	Q9lar2 rickettsia
318	6	1.4	196	2	Q9LAQ9	Q9laq9 rickettsia
319	6	1.4	196	2	Q9LAQ8	Q9laq8 rickettsia
320	6	1.4	197	2	Q48372	Q48372 israeli tic
321	6	1.4	197	2	Q52645	Q52645 rickettsia
322	6	1.4	197	2	Q52655	Q52655 rickettsia
323	6	1.4	197	2	Q52669	Q52669 rickettsia
324	6	1.4	197	2	Q52758	Q52758 rickettsia
325	6	1.4	197	2	Q52765	Q52765 rickettsia
326	6	1.4	197	2	Q52895	Q52895 rickettsia
327	6	1.4	197	2	Q53006	Q53006 rickettsia
328	6	1.4	197	2	Q53024	Q53024 rickettsia
329	6	1.4	197	2	Q53108	Q53108 rickettsia
330	6	1.4	197	2	Q53109	Q53109 rickettsia
331	6	1.4	197	2	Q53110	Q53110 rickettsia
332	6	1.4	197	2	Q57548	Q57548 rickettsia
333	6	1.4	197	2	Q9LBK9	Q9lbk9 gluconobact
334	6	1.4	197	2	Q44180	Q44180 astrakhan r
335	6	1.4	197	2	Q9HU41	Q9hu41 pseudomonas
336	6	1.4	197	2	Q56384	Q56384 rickettsia
337	6	1.4	198	2	Q52667	Q52667 rickettsia
338	6	1.4	198	2	Q9RFV7	Q9rfv7 vibrio harv
339	6	1.4	200	14	Q9W8K2	Q9w8k2 human herpe
340	6	1.4	200	14	Q9PX32	Q9px32 human herpe
341	6	1.4	200	14	Q9J5E7	Q9j5e7 fowlpox vir
342	6	1.4	201	4	Q9NU14	Q9nu14 homo sapien
343	6	1.4	202	5	Q21905	Q21905 caenorhabdi
344	6	1.4	202	10	O65136	O65136 syringa vul
345	6	1.4	203	2	Q9R6K2	Q9r6k2 agrobacteri
346	6	1.4	203	2	Q9HTW2	Q9htw2 pseudomonas
347	6	1.4	204	2	Q9FDP2	Q9fdp2 rickettsia
348	6	1.4	204	2	Q9F9R6	Q9f9r6 rickettsia
349	6	1.4	204	2	Q9F9R4	Q9f9r4 rickettsia
350	6	1.4	206	1	Q9HSV9	Q9hsv9 halobacteri
351	6	1.4	210	2	Q9KJH3	Q9kjh3 rickettsia
352	6	1.4	210	4	Q9UE79	Q9ue79 homo sapien
353	6	1.4	212	8	O79625	O79625 russula ros
354	6	1.4	213	2	Q56417	Q56417 thermus aqu
355	6	1.4	213	8	O79623	O79623 gomphus flo
356	6	1.4	214	1	Q9HL63	Q9hl63 thermoplasm
357	6	1.4	214	2	Q57013	Q57013 borrelia lo
358	6	1.4	216	1	O30209	O30209 archaeoglob
359	6	1.4	218	2	P73898	P73898 synechocyst
360	6	1.4	219	8	Q02669	Q02669 podospora a
361	6	1.4	220	10	Q42500	Q42500 rumex aceto
362	6	1.4	221	5	Q9V5H1	Q9v5h1 drosophila
363	6	1.4	226	2	Q9EYN9	Q9eyn9 synechococc
364	6	1.4	226	5	Q9V360	Q9v360 drosophila
365	6	1.4	228	8	Q9XNB2	Q9xnb2 photinus ob
366	6	1.4	230	5	Q9VAK2	Q9vak2 drosophila
367	6	1.4	232	2	Q60011	Q60011 streptomyce
368	6	1.4	232	10	Q43507	Q43507 lycopersico
369	6	1.4	233	1	Q58491	Q58491 methanococc
370	6	1.4	233	5	O96361	O96361 hyphantria
371	6	1.4	236	2	Q9ZH58	Q9zh58 streptomyce
372	6	1.4	236	3	Q12282	Q12282 saccharomyc
373	6	1.4	236	10	Q9FZK8	Q9fzk8 arabidopsis
374	6	1.4	237	2	Q9K8U2	Q9k8u2 bacillus ha
375	6	1.4	241	9	Q9T1E5	Q9t1e5 lactobacill
376	6	1.4	241	10	Q9FZK9	Q9fzk9 arabidopsis
377	6	1.4	242	2	Q49662	Q49662 mycobacteri
378	6	1.4	244	4	Q9NPX1	Q9npx1 homo sapien
379	6	1.4	244	10	Q9LWC5	Q9lwc5 oryza sativ
380	6	1.4	247	2	O30928	O30928 streptococc
381	6	1.4	248	2	Q9PMY5	Q9pmy5 campylobact
382	6	1.4	249	2	P72275	P72275 rhizobium f
383	6	1.4	252	2	O05162	O05162 synechococc
384	6	1.4	252	10	O48692	O48692 arabidopsis

385	6	1.4	253	2	Q9L4H1	Q9l4h1 streptococc
386	6	1.4	253	5	Q9N3G0	Q9n3g0 caenorhabdi
387	6	1.4	254	2	Q9PCC3	Q9pcc3 xylella fas
388	6	1.4	255	2	Q9RWX1	Q9rwx1 deinococcus
389	6	1.4	255	2	Q9PGC8	Q9pgc8 xylella fas
390	6	1.4	257	7	O19430	O19430 felis silve
391	6	1.4	258	2	O07601	O07601 bacillus su
392	6	1.4	258	2	Q9X4H5	Q9x4h5 streptomyce
393	6	1.4	260	2	O31229	O31229 arthrobacte
394	6	1.4	260	2	Q9KG24	Q9kg24 bacillus ha
395	6	1.4	262	5	Q9NIM7	Q9nim7 babesia mic
396	6	1.4	264	2	Q9PGT5	Q9pgt5 xylella fas
397	6	1.4	264	5	Q9UB64	Q9ub64 drosophila
398	6	1.4	264	5	Q9UB53	Q9ub53 drosophila
399	6	1.4	264	5	Q9TW58	Q9tw58 drosophila
400	6	1.4	265	2	Q9PPN9	Q9ppn9 ureaplasma
401	6	1.4	266	4	Q14666	Q14666 homo sapien
402	6	1.4	267	2	Q9RU14	Q9ru14 deinococcus
403	6	1.4	267	2	Q9K912	Q9k912 bacillus ha
404	6	1.4	267	4	Q9NUT4	Q9nut4 homo sapien
405	6	1.4	268	5	Q9V4A9	Q9v4a9 drosophila
406	6	1.4	269	3	O94212	O94212 gigaspora m
407	6	1.4	269	5	Q9NCE6	Q9nce6 mastotermes
408	6	1.4	269	8	Q9ZZV9	Q9zzv9 saccharomyc
409	6	1.4	270	4	Q9H6N3	Q9h6n3 homo sapien
410	6	1.4	270	5	Q9NCE5	Q9nce5 mastotermes
411	6	1.4	271	14	Q69374	Q69374 gallid herp
412	6	1.4	272	2	Q51013	Q51013 neisseria g
413	6	1.4	272	10	Q9ZTA9	Q9zta9 dolichos la
414	6	1.4	272	10	Q9FHA1	Q9fha1 arabidopsis
415	6	1.4	273	2	O25281	O25281 helicobacte
416	6	1.4	273	3	O94670	O94670 schizosacch
417	6	1.4	274	1	O28585	O28585 archaeoglob
418	6	1.4	274	5	Q9N4G7	Q9n4g7 caenorhabdi
419	6	1.4	277	1	O58047	O58047 pyrococcus
420	6	1.4	277	1	Q9UZU8	Q9uzu8 pyrococcus
421	6	1.4	277	2	O54371	O54371 treponema h
422	6	1.4	277	2	Q9XD60	Q9xd60 pseudomonas
423	6	1.4	278	2	O06991	O06991 bacillus su
424	6	1.4	278	2	Q9Z9J3	Q9z9j3 bacillus ha
425	6	1.4	278	2	Q9KB62	Q9kb62 bacillus ha
426	6	1.4	279	2	Q9K6M6	Q9k6m6 bacillus ha
427	6	1.4	279	5	Q9TYE1	Q9tye1 plasmodium
428	6	1.4	280	6	O97624	O97624 canis famil
429	6	1.4	281	13	Q9DG09	Q9dg09 gallus gall
430	6	1.4	283	2	O31397	O31397 bacillus st
431	6	1.4	283	13	O93417	O93417 gallus gall
432	6	1.4	284	2	Q9RU83	Q9ru83 deinococcus
433	6	1.4	285	5	Q9U632	Q9u632 trichomegal
434	6	1.4	287	10	Q38818	Q38818 arabidopsis
435	6	1.4	288	14	Q9EME3	Q9eme3 amsacta moo
436	6	1.4	288	14	Q9DKY5	Q9dky5 turnip mosa
437	6	1.4	289	5	Q9NIN4	Q9nin4 babesia mic
438	6	1.4	291	5	Q9VI51	Q9vi51 drosophila
439	6	1.4	291	10	Q9SQ78	Q9sq78 arabidopsis
440	6	1.4	291	11	Q9JMD3	Q9jmd3 mus musculu
441	6	1.4	292	2	O25957	O25957 helicobacte
442	6	1.4	292	2	Q9ZJK7	Q9zjk7 helicobacte
443	6	1.4	292	2	Q9KKQ3	Q9kkq3 vibrio chol
444	6	1.4	294	10	Q9SWZ8	Q9swz8 arabidopsis
445	6	1.4	295	2	Q46990	Q46990 enterobacte
446	6	1.4	295	2	Q9KXQ7	Q9kxq7 streptomyce
447	6	1.4	298	5	Q9NIN5	Q9nin5 babesia mic
448	6	1.4	299	4	Q9NYW7	Q9nyw7 homo sapien
449	6	1.4	299	5	P91994	P91994 caenorhabdi
450	6	1.4	301	2	O67352	O67352 aquifex aeo
451	6	1.4	301	5	Q9GUQ7	Q9guq7 leishmania
452	6	1.4	301	10	Q9SMQ1	Q9smq1 arabidopsis
453	6	1.4	302	1	Q9HP57	Q9hp57 halobacteri
454	6	1.4	302	2	P96939	P96939 mycobacteri
455	6	1.4	303	2	Q9S1K9	Q9s1k9 streptovert
456	6	1.4	303	2	Q9KCH1	Q9kch1 bacillus ha
457	6	1.4	305	10	Q9ZVR5	Q9zvr5 arabidopsis
458	6	1.4	306	10	O23551	O23551 arabidopsis
459	6	1.4	308	10	Q9LGP8	Q9lgp8 oryza sativ
460	6	1.4	308	14	Q9IR22	Q9ir22 dianthoviru
461	6	1.4	310	10	Q9SJZ1	Q9sjz1 arabidopsis
462	6	1.4	310	10	Q9M4Q1	Q9m4q1 brassica ra
463	6	1.4	312	1	Q9HJM7	Q9hjm7 thermoplasm
464	6	1.4	313	2	P97213	P97213 clostridium
465	6	1.4	313	5	O97363	O97363 bombyx mori
466	6	1.4	314	2	Q9X9X3	Q9x9x3 streptomyce
467	6	1.4	315	5	Q9U2L8	Q9u2l8 caenorhabdi
468	6	1.4	316	5	Q23077	Q23077 caenorhabdi
469	6	1.4	317	2	Q9KLB2	Q9klb2 vibrio chol
470	6	1.4	317	5	Q18687	Q18687 caenorhabdi
471	6	1.4	317	5	Q23594	Q23594 caenorhabdi
472	6	1.4	317	14	Q9J849	Q9j849 spodoptera
473	6	1.4	318	1	O58858	O58858 pyrococcus
474	6	1.4	319	2	Q9KZK9	Q9kzk9 streptomyce
475	6	1.4	319	10	Q40667	Q40667 oryza sativ
476	6	1.4	319	14	Q9WBV0	Q9wbv0 influenza a
477	6	1.4	319	14	Q9WBU9	Q9wbu9 influenza a
478	6	1.4	319	14	Q9WBU7	Q9wbu7 influenza a
479	6	1.4	319	14	Q9W8Y7	Q9w8y7 influenza a
480	6	1.4	320	5	O16400	O16400 caenorhabdi
481	6	1.4	320	10	Q9LWF1	Q9lwf1 oryza sativ
482	6	1.4	321	4	Q9H6E8	Q9h6e8 homo sapien
483	6	1.4	321	5	O01538	O01538 caenorhabdi
484	6	1.4	322	4	Q14968	Q14968 homo sapien
485	6	1.4	322	8	Q9T6Q7	Q9t6q7 trachemys s
486	6	1.4	323	2	Q9RVU2	Q9rvu2 deinococcus
487	6	1.4	323	5	Q19121	Q19121 caenorhabdi
488	6	1.4	323	10	Q42993	Q42993 oryza sativ
489	6	1.4	324	2	Q50549	Q50549 mycobacteri
490	6	1.4	324	4	Q9Y310	Q9y310 homo sapien
491	6	1.4	324	5	Q9Y0D2	Q9y0d2 hypera post
492	6	1.4	324	14	O37399	O37399 chlorella v
493	6	1.4	325	5	Q21104	Q21104 caenorhabdi
494	6	1.4	326	2	O86279	O86279 lactococcus
495	6	1.4	326	5	Q9NIP2	Q9nip2 babesia mic
496	6	1.4	327	8	Q9MUR4	Q9mur4 mesostigma
497	6	1.4	327	14	Q9WBV4	Q9wbv4 influenza a
498	6	1.4	327	14	Q9WBV2	Q9wbv2 influenza a
499	6	1.4	327	14	Q9WBV1	Q9wbv1 influenza a
500	6	1.4	327	14	Q9WBU8	Q9wbu8 influenza a
501	6	1.4	327	14	Q9WBU6	Q9wbu6 influenza a
502	6	1.4	327	14	Q9WBU5	Q9wbu5 influenza a
503	6	1.4	327	14	Q9WBU4	Q9wbu4 influenza a
504	6	1.4	327	14	Q9WBU3	Q9wbu3 influenza a
505	6	1.4	327	14	Q9WBU2	Q9wbu2 influenza a
506	6	1.4	327	14	Q9WBU1	Q9wbu1 influenza a
507	6	1.4	327	14	Q9WBU0	Q9wbu0 influenza a
508	6	1.4	327	14	Q9WBT9	Q9wbt9 influenza a
509	6	1.4	327	14	Q9WBT8	Q9wbt8 influenza a
510	6	1.4	328	2	Q9KRN4	Q9krn4 vibrio chol
511	6	1.4	328	13	Q9PSJ3	Q9psj3 ictalurus p
512	6	1.4	328	14	Q9WBV3	Q9wbv3 influenza a
513	6	1.4	329	2	O07506	O07506 bacillus me
514	6	1.4	330	3	Q9P7P7	Q9p7p7 schizosacch
515	6	1.4	331	2	Q9RIX8	Q9rix8 streptomyce
516	6	1.4	331	2	Q9F513	Q9f513 escherichia
517	6	1.4	332	2	O84491	O84491 chlamydia t
518	6	1.4	332	5	Q9NAN0	Q9nan0 caenorhabdi
519	6	1.4	333	5	Q9Y071	Q9y071 periplaneta
520	6	1.4	335	3	O74235	O74235 schizosacch
521	6	1.4	335	3	Q9P7I2	Q9p7i2 schizosacch
522	6	1.4	335	6	Q9GM42	Q9gm42 macaca fasc
523	6	1.4	336	2	Q45946	Q45946 coxiella bu
524	6	1.4	337	5	Q9GSD9	Q9gsd9 strongyloce
525	6	1.4	338	1	Q9V1N6	Q9v1n6 pyrococcus
526	6	1.4	338	3	O74426	O74426 schizosacch
527	6	1.4	338	10	O81017	O81017 arabidopsis
528	6	1.4	340	10	O23800	O23800 ananas como
529	6	1.4	340	14	O57262	O57262 vaccinia vi
530	6	1.4	340	14	O72752	O72752 cowpox viru

531	6	1.4	341	2	Q9HVG0	Q9hvg0 pseudomonas
532	6	1.4	341	3	O74509	O74509 schizosacch
533	6	1.4	342	5	Q22209	Q22209 caenorhabdi
534	6	1.4	343	1	O26922	O26922 methanobact
535	6	1.4	344	2	Q9z950	Q9z950 chlamydia p
536	6	1.4	344	5	O45603	O45603 caenorhabdi
537	6	1.4	345	3	Q06494	Q06494 saccharomyc
538	6	1.4	346	2	Q9K225	Q9k225 chlamydia p
539	6	1.4	346	5	Q9U2L5	Q9u215 caenorhabdi
540	6	1.4	346	8	Q9ML71	Q9ml71 coturnix au
541	6	1.4	347	2	Q59797	Q59797 neisseria g
542	6	1.4	347	5	Q18732	Q18732 caenorhabdi
543	6	1.4	348	3	O74789	O74789 schizosacch
544	6	1.4	350	5	Q17586	Q17586 caenorhabdi
545	6	1.4	351	2	O34024	O34024 chlamydia p
546	6	1.4	352	10	O24641	O24641 ananas como
547	6	1.4	353	2	Q9KL63	Q9kl63 vibrio chol
548	6	1.4	353	5	Q9NB25	Q9nb25 heliothis z
549	6	1.4	354	5	O44458	O44458 caenorhabdi
550	6	1.4	354	14	Q9IP69	Q9ip69 human cytom
551	6	1.4	354	14	Q9IP68	Q9ip68 human cytom
552	6	1.4	354	14	Q9IP67	Q9ip67 human cytom
553	6	1.4	354	14	Q9IP66	Q9ip66 human cytom
554	6	1.4	354	14	Q9IP65	Q9ip65 human cytom
555	6	1.4	355	1	O27789	O27789 methanobact
556	6	1.4	356	5	Q18434	Q18434 caenorhabdi
557	6	1.4	356	14	Q82515	Q82515 influenza a
558	6	1.4	357	2	Q9S1Q2	Q9s1q2 streptomyce
559	6	1.4	357	5	Q19776	Q19776 caenorhabdi
560	6	1.4	358	2	Q9K1B4	Q9k1b4 neisseria m
561	6	1.4	358	2	Q9JX88	Q9jx88 neisseria m
562	6	1.4	359	4	Q9Y365	Q9y365 homo sapien
563	6	1.4	360	3	Q9P8B3	Q9p8b3 agaricus bi
564	6	1.4	360	5	Q19549	Q19549 caenorhabdi
565	6	1.4	361	3	Q9P5Y2	Q9p5y2 neurospora
566	6	1.4	361	5	Q9U235	Q9u235 caenorhabdi
567	6	1.4	362	5	Q9NIP1	Q9nip1 babesia mic
568	6	1.4	362	5	Q9V817	Q9v817 drosophila
569	6	1.4	362	14	Q82517	Q82517 influenza a
570	6	1.4	363	10	Q9XFV3	Q9xfv3 chlamydomon
571	6	1.4	364	5	Q9V9X4	Q9v9x4 drosophila
572	6	1.4	365	2	Q9S247	Q9s247 streptomyce
573	6	1.4	369	1	Q9HT02	Q9ht02 halobacteri
574	6	1.4	370	2	Q9F903	Q9f903 streptococc
575	6	1.4	370	10	Q9M1J6	Q9mlj6 arabidopsis
576	6	1.4	371	10	Q9SZH2	Q9szh2 arabidopsis
577	6	1.4	377	2	O51302	O51302 borrelia bu
578	6	1.4	377	14	Q70662	Q70662 human immun
579	6	1.4	378	14	Q9WRR2	Q9wrr2 macaca mula
580	6	1.4	379	6	Q9GL13	Q9gl13 macropus ru
581	6	1.4	380	5	Q9VZX8	Q9vzx8 drosophila
582	6	1.4	381	9	Q38310	Q38310 lactococcus
583	6	1.4	381	10	Q9LF94	Q9lf94 arabidopsis
584	6	1.4	382	1	Q9HLJ8	Q9hlj8 thermoplasm
585	6	1.4	383	1	O27530	O27530 methanobact
586	6	1.4	383	6	Q9GL20	Q9gl20 cynopterus
587	6	1.4	384	3	O14094	O14094 schizosacch
588	6	1.4	385	4	Q9H8X7	Q9h8x7 homo sapien
589	6	1.4	385	6	Q9GL14	Q9gl14 micropotamo
590	6	1.4	385	6	Q9GKZ7	Q9gkz7 tenrec ecau
591	6	1.4	386	2	Q9XDK1	Q9xdk1 bacteroides
592	6	1.4	386	3	O13901	O13901 schizosacch
593	6	1.4	387	4	Q9UKT5	Q9ukt5 homo sapien
594	6	1.4	387	6	Q9GL17	Q9gl17 felis silve
595	6	1.4	388	6	Q9GL18	Q9gl18 diceros bic
596	6	1.4	388	9	Q38198	Q38198 xanthomonas
597	6	1.4	389	6	Q9GL07	Q9gl07 phoca vitul
598	6	1.4	389	6	Q9GKZ5	Q9gkz5 tupaia tana
599	6	1.4	390	3	Q9P3B6	Q9p3b6 neurospora
600	6	1.4	390	6	Q9GL19	Q9gl19 cynocephalu
601	6	1.4	390	6	Q9GKZ6	Q9gkz6 trichechus
602	6	1.4	391	2	Q9FD61	Q9fd61 streptococc
603	6	1.4	391	6	Q9GL06	Q9gl06 sus scrofa
604	6	1.4	392	5	Q9VHE6	Q9vhe6 drosophila
605	6	1.4	392	6	Q9GL35	Q9gl35 balaenopter
606	6	1.4	393	6	Q9GL16	Q9gl16 hippopotamu
607	6	1.4	393	6	Q9GL15	Q9gl15 macrotus ca
608	6	1.4	393	6	Q9GL11	Q9gl11 nycticebus
609	6	1.4	393	11	Q9JJW2	Q9jjw2 cavia porce
610	6	1.4	395	2	Q9WZJ6	Q9wzj6 thermotoga
611	6	1.4	395	6	Q9GL12	Q9gl12 manis sp. a
612	6	1.4	395	10	Q9ZT98	Q9zt98 arabidopsis
613	6	1.4	395	10	Q9LR79	Q9lr79 arabidopsis
614	6	1.4	396	2	O54055	O54055 ruminococcu
615	6	1.4	397	1	Q58479	Q58479 methanococc
616	6	1.4	398	4	Q9NUJ0	Q9nuj0 homo sapien
617	6	1.4	399	1	Q9UXX6	Q9uxx6 pyrococcus
618	6	1.4	401	5	Q9VQ24	Q9vq24 drosophila
619	6	1.4	402	10	Q9FJ89	Q9fj89 arabidopsis
620	6	1.4	403	1	O29552	O29552 archaeoglob
621	6	1.4	403	14	Q9YVI7	Q9yvi7 melanoplus
622	6	1.4	404	5	Q9XVC5	Q9xvc5 caenorhabdi
623	6	1.4	406	2	Q9L447	Q9l447 streptococc
624	6	1.4	406	2	Q9L389	Q9l389 streptococc
625	6	1.4	406	2	Q9EZP9	Q9ezp9 streptococc
626	6	1.4	406	2	Q9EZP8	Q9ezp8 streptococc
627	6	1.4	406	2	Q9EZP7	Q9ezp7 streptococc
628	6	1.4	406	2	Q9EZP6	Q9ezp6 streptococc
629	6	1.4	406	10	Q9ZWJ6	Q9zwj6 zea mays (m
630	6	1.4	407	2	Q9RK30	Q9rk30 streptomyce
631	6	1.4	408	8	Q9XP23	Q9xp23 marmota men
632	6	1.4	410	10	Q9FT97	Q9ft97 arabidopsis
633	6	1.4	411	2	O66182	O66182 arthrobacte
634	6	1.4	411	2	Q9KS65	Q9ks65 vibrio chol
635	6	1.4	412	8	Q9G3Z1	Q9g3z1 cottus bair
636	6	1.4	412	8	Q9G3Y8	Q9g3y8 cottus caro
637	6	1.4	412	8	Q9G3Y5	Q9g3y5 cottus hyps
638	6	1.4	414	2	O66163	O66163 vibrio para
639	6	1.4	414	4	Q9UKW8	Q9ukw8 homo sapien
640	6	1.4	414	4	Q9NTJ9	Q9ntj9 homo sapien
641	6	1.4	414	14	O89241	O89241 gallid herp
642	6	1.4	415	14	Q66626	Q66626 equine herp
643	6	1.4	417	4	Q9NT19	Q9nt19 homo sapien
644	6	1.4	418	10	Q9SCD6	Q9scd6 fagus sylva
645	6	1.4	419	10	Q9ZSL6	Q9zsl6 arabidopsis
646	6	1.4	422	10	Q9FI34	Q9fi34 arabidopsis
647	6	1.4	422	14	Q89034	Q89034 vesicular s
648	6	1.4	422	14	Q89035	Q89035 vesicular s
649	6	1.4	422	14	Q89036	Q89036 vesicular s
650	6	1.4	422	14	Q89037	Q89037 vesicular s
651	6	1.4	422	14	Q89038	Q89038 vesicular s
652	6	1.4	422	14	Q89039	Q89039 vesicular s
653	6	1.4	422	14	Q89040	Q89040 vesicular s
654	6	1.4	422	14	Q89594	Q89594 vesicular s
655	6	1.4	422	14	Q89779	Q89779 vesicular s
656	6	1.4	423	5	Q21865	Q21865 caenorhabdi
657	6	1.4	423	13	O93256	O93256 gallus gall
658	6	1.4	424	8	Q9THI5	Q9thi5 epistephium
659	6	1.4	425	10	Q9FME6	Q9fme6 arabidopsis
660	6	1.4	427	8	Q9GEA8	Q9gea8 hypopterygi
661	6	1.4	428	1	Q9UYV9	Q9uyv9 pyrococcus
662	6	1.4	429	10	O64511	O64511 arabidopsis
663	6	1.4	429	13	Q9PWD8	Q9pwd8 brachydanio
664	6	1.4	430	14	Q9YVT5	Q9yvt5 melanoplus
665	6	1.4	432	10	Q9M884	Q9m884 arabidopsis
666	6	1.4	433	11	Q9QWL7	Q9qwl7 mus musculu
667	6	1.4	434	2	Q9WYG5	Q9wyg5 thermotoga
668	6	1.4	434	10	Q9LFM7	Q9lfm7 arabidopsis
669	6	1.4	434	13	Q9IB76	Q9ib76 pagrus majo
670	6	1.4	436	10	Q39073	Q39073 arabidopsis
671	6	1.4	438	2	Q9PHP8	Q9php8 campylobact
672	6	1.4	438	8	O80030	O80030 epistephium
673	6	1.4	439	2	O05495	O05495 bacillus su
674	6	1.4	440	5	O96686	O96686 drosophila
675	6	1.4	440	5	Q9VJ02	Q9vj02 drosophila
676	6	1.4	440	8	O78589	O78589 epistephium

677	6	1.4	440	8	Q9MFK0	Q9mfk0 diabrotica
678	6	1.4	440	10	Q9LER8	Q9ler8 arabidopsis
679	6	1.4	440	10	Q41644	Q41644 volvox cart
680	6	1.4	440	13	O42273	O42273 gallus gall
681	6	1.4	441	8	Q37144	Q37144 apostasia s
682	6	1.4	441	10	O22282	O22282 arabidopsis
683	6	1.4	441	10	Q9FJT1	Q9fjt1 arabidopsis
684	6	1.4	442	2	O07845	O07845 burkholderi
685	6	1.4	444	10	Q9FW55	Q9fw55 oryza sativ
686	6	1.4	448	5	Q9VHE3	Q9vhe3 drosophila
687	6	1.4	448	8	O78587	O78587 epistephium
688	6	1.4	449	5	Q9XZN7	Q9xzn7 asterias ru
689	6	1.4	449	8	O78588	O78588 epistephium
690	6	1.4	449	10	Q9FM55	Q9fm55 arabidopsis
691	6	1.4	450	2	O50511	O50511 streptomyce
692	6	1.4	450	3	O74829	O74829 schizosacch
693	6	1.4	451	8	Q9T7K5	Q9t7k5 crassostrea
694	6	1.4	455	5	Q9GT60	Q9gt60 drosophila
695	6	1.4	455	5	Q9GT59	Q9gt59 drosophila
696	6	1.4	455	5	Q9GT58	Q9gt58 drosophila
697	6	1.4	455	5	Q9GT57	Q9gt57 drosophila
698	6	1.4	455	5	Q9GNE8	Q9gne8 drosophila
699	6	1.4	455	5	Q9GND6	Q9gnd6 drosophila
700	6	1.4	455	5	Q9GN78	Q9gn78 drosophila
701	6	1.4	455	5	Q9GN21	Q9gn21 drosophila
702	6	1.4	455	13	Q07426	Q07426 carassius a
703	6	1.4	456	2	Q9KHI4	Q9khi4 haemophilus
704	6	1.4	456	2	Q9KHI3	Q9khi3 haemophilus
705	6	1.4	456	2	Q9KHI2	Q9khi2 haemophilus
706	6	1.4	456	5	Q9TY15	Q9ty15 nyctotherus
707	6	1.4	456	5	Q9TX28	Q9tx28 lytechinus
708	6	1.4	458	3	Q12104	Q12104 saccharomyc
709	6	1.4	458	5	Q9W1D1	Q9w1d1 drosophila
710	6	1.4	458	8	Q9TA03	Q9ta03 lampetra fl
711	6	1.4	459	4	Q9H673	Q9h673 homo sapien
712	6	1.4	460	8	O47759	O47759 xenocypris
713	6	1.4	460	8	O47760	O47760 distoechodo
714	6	1.4	460	8	O47761	O47761 distoechodo
715	6	1.4	460	8	O47763	O47763 myxocyprinu
716	6	1.4	460	8	O47764	O47764 xenocypris
717	6	1.4	460	8	O47766	O47766 pseudobrama
718	6	1.4	460	8	O47767	O47767 xenocypris
719	6	1.4	460	8	O47768	O47768 xenocypris
720	6	1.4	460	8	O47770	O47770 xenocypris
721	6	1.4	460	8	O47771	O47771 xenocyprioi
722	6	1.4	460	8	O47772	O47772 xenocypris
723	6	1.4	460	8	O47774	O47774 xenocypris
724	6	1.4	460	8	O47776	O47776 xenocypris
725	6	1.4	460	8	O47777	O47777 xenocypris
726	6	1.4	460	8	O47778	O47778 xenocypris
727	6	1.4	460	8	O47780	O47780 xenocyprioi
728	6	1.4	460	8	O47782	O47782 xenocypris
729	6	1.4	460	8	Q34046	Q34046 cyprinus ca
730	6	1.4	460	8	Q9XKX4	Q9xkx4 chrysemys p
731	6	1.4	460	8	Q9T918	Q9t918 aristichthy
732	6	1.4	460	8	Q9T917	Q9t917 cultrichthy
733	6	1.4	460	8	Q9T915	Q9t915 abbottina r
734	6	1.4	460	8	Q9T914	Q9t914 ctenopharyn
735	6	1.4	460	8	Q9T913	Q9t913 hypophthalm
736	6	1.4	460	8	Q9T912	Q9t912 megalobrama
737	6	1.4	460	8	Q9T910	Q9t910 megalobrama
738	6	1.4	460	8	Q9T909	Q9t909 megalobrama
739	6	1.4	460	8	Q9T908	Q9t908 megalobrama
740	6	1.4	460	8	Q9T907	Q9t907 parabramis
741	6	1.4	460	8	Q9T905	Q9t905 sinocycloch
742	6	1.4	460	8	Q9T903	Q9t903 schizothora
743	6	1.4	460	8	Q9T902	Q9t902 semilabeo p
744	6	1.4	460	8	Q9G6S8	Q9g6s8 sardinops m
745	6	1.4	460	8	Q9G6R6	Q9g6r6 coregonus l
746	6	1.4	460	8	Q9G6Q4	Q9g6q4 diplophos t
747	6	1.4	460	8	Q9G6P1	Q9g6p1 polymixia j
748	6	1.4	460	11	Q62739	Q62739 rattus norv
749	6	1.4	461	8	Q9XN32	Q9xn32 salvelinus
750	6	1.4	461	8	Q9XN24	Q9xn24 salvelinus
751	6	1.4	462	4	Q9ULK6	Q9ulk6 homo sapien
752	6	1.4	463	2	P74136	P74136 synechocyst
753	6	1.4	464	5	O02068	O02068 caenorhabdi
754	6	1.4	464	5	Q9GTP3	Q9gtp3 entamoeba h
755	6	1.4	465	5	O02206	O02206 caenorhabdi
756	6	1.4	466	4	Q9NUS2	Q9nus2 homo sapien
757	6	1.4	466	8	P93893	P93893 burchardia
758	6	1.4	467	1	Q9HH54	Q9hh54 methanobact
759	6	1.4	467	10	Q9SY75	Q9sy75 arabidopsis
760	6	1.4	468	5	Q9GTX0	Q9gtx0 entamoeba h
761	6	1.4	468	9	O80199	O80199 methanobact
762	6	1.4	470	2	O06546	O06546 mycobacteri
763	6	1.4	470	10	Q9LLE0	Q9lle0 solanum tub
764	6	1.4	470	10	Q9FN78	Q9fn78 arabidopsis
765	6	1.4	471	2	Q9HWN7	Q9hwn7 pseudomonas
766	6	1.4	472	1	Q9UYK4	Q9uyk4 pyrococcus
767	6	1.4	472	10	Q9FXZ7	Q9fxz7 zea mays (m
768	6	1.4	473	4	Q9H0Q5	Q9h0q5 homo sapien
769	6	1.4	473	10	Q9SHZ2	Q9shz2 arabidopsis
770	6	1.4	473	10	Q9FTN0	Q9ftn0 oryza sativ
771	6	1.4	473	10	Q9FLX9	Q9flx9 arabidopsis
772	6	1.4	474	5	Q9VTR1	Q9vtr1 drosophila
773	6	1.4	475	4	O43818	O43818 homo sapien
774	6	1.4	475	8	Q9MTF1	Q9mtf1 tacca leont
775	6	1.4	476	8	O98728	O98728 tacca chant
776	6	1.4	477	2	Q9I6R8	Q9i6r8 pseudomonas
777	6	1.4	477	4	Q9P1Y7	Q9ply7 homo sapien
778	6	1.4	477	5	Q18153	Q18153 caenorhabdi
779	6	1.4	478	3	Q06001	Q06001 saccharomyc
780	6	1.4	479	2	Q9KUW8	Q9kuw8 vibrio chol
781	6	1.4	479	10	Q9LY28	Q9ly28 arabidopsis
782	6	1.4	481	1	O58634	O58634 pyrococcus
783	6	1.4	481	2	Q9FC47	Q9fc47 streptomyce
784	6	1.4	481	10	Q9LYD5	Q9lyd5 arabidopsis
785	6	1.4	483	5	Q19463	Q19463 caenorhabdi
786	6	1.4	483	8	Q9G922	Q9g922 ochromonas
787	6	1.4	483	14	Q9WRJ1	Q9wrj1 human immun
788	6	1.4	484	8	O47458	O47458 crypthecodi
789	6	1.4	485	5	Q9V9N1	Q9v9n1 drosophila
790	6	1.4	487	2	Q9KQQ3	Q9kqq3 vibrio chol
791	6	1.4	487	10	Q9FI43	Q9fi43 arabidopsis
792	6	1.4	488	10	Q9SK58	Q9sk58 arabidopsis
793	6	1.4	490	2	O84667	O84667 chlamydia t
794	6	1.4	490	2	Q9PLR4	Q9plr4 chlamydia m
795	6	1.4	490	2	Q9K5Q0	Q9k5q0 bacillus ha
796	6	1.4	490	5	O17655	O17655 caenorhabdi
797	6	1.4	491	5	Q9V7C6	Q9v7c6 drosophila
798	6	1.4	491	10	O22714	O22714 arabidopsis
799	6	1.4	494	2	Q9F1M9	Q9f1m9 shewanella
800	6	1.4	494	5	O02641	O02641 caenorhabdi
801	6	1.4	494	8	Q9G860	Q9g860 malawimonas
802	6	1.4	495	4	Q9Y5L3	Q9y5l3 homo sapien
803	6	1.4	495	11	Q9JHY5	Q9jhy5 rattus norv
804	6	1.4	495	11	O55026	O55026 mus musculu
805	6	1.4	495	11	O35795	O35795 rattus norv
806	6	1.4	496	2	Q9RS47	Q9rs47 deinococcus
807	6	1.4	496	10	Q9LZZ0	Q9lzz0 arabidopsis
808	6	1.4	496	10	Q9FM80	Q9fm80 arabidopsis
809	6	1.4	497	5	Q9XYV9	Q9xyv9 drosophila
810	6	1.4	497	5	Q9VVX5	Q9vvx5 drosophila
811	6	1.4	498	1	O27314	O27314 methanobact
812	6	1.4	498	1	Q9V2Y5	Q9v2y5 methanobact
813	6	1.4	498	2	Q9KBS6	Q9kbs6 bacillus ha
814	6	1.4	499	10	O49540	O49540 arabidopsis
815	6	1.4	500	6	Q29454	Q29454 bos taurus
816	6	1.4	500	10	O22926	O22926 arabidopsis
817	6	1.4	501	5	Q9W133	Q9w133 drosophila
818	6	1.4	504	2	Q59558	Q59558 mycobacteri
819	6	1.4	504	8	Q9G919	Q9g919 ochromonas
820	6	1.4	504	11	Q63092	Q63092 rattus norv
821	6	1.4	506	4	Q9H625	Q9h625 homo sapien
822	6	1.4	506	10	Q9LPL6	Q9lpl6 arabidopsis

823	6	1.4	507	3	Q9HGR6	Q9hgr6 mucor circi
824	6	1.4	507	5	Q9N3F2	Q9n3f2 caenorhabdi
825	6	1.4	508	8	Q9TKP8	Q9tkp8 lotus pursh
826	6	1.4	508	11	O54896	O54896 mus musculu
827	6	1.4	509	2	Q9X6P3	Q9x6p3 neisseria g
828	6	1.4	512	2	O08251	O08251 rhizobium m
829	6	1.4	512	8	Q9XMS7	Q9xms7 tetrahymena
830	6	1.4	513	5	Q93690	Q93690 caenorhabdi
831	6	1.4	513	8	Q9TKQ0	Q9tkq0 ornithopus
832	6	1.4	514	10	O24224	O24224 oryza sativ
833	6	1.4	515	2	Q9RYW9	Q9ryw9 deinococcus
834	6	1.4	517	5	Q21292	Q21292 caenorhabdi
835	6	1.4	518	2	O84201	O84201 chlamydia t
836	6	1.4	519	3	O60163	O60163 schizosacch
837	6	1.4	520	5	Q25643	Q25643 plasmodium
838	6	1.4	520	5	Q21385	Q21385 caenorhabdi
839	6	1.4	522	2	Q9RGE6	Q9rge6 clostridium
840	6	1.4	523	5	Q18236	Q18236 caenorhabdi
841	6	1.4	523	10	Q9LVF9	Q9lvf9 arabidopsis
842	6	1.4	524	11	Q9R0H5	Q9r0h5 mus musculu
843	6	1.4	525	2	O83500	O83500 treponema p
844	6	1.4	526	1	Q9HMX1	Q9hmx1 halobacteri
845	6	1.4	528	5	Q9VR38	Q9vr38 drosophila
846	6	1.4	529	10	Q9FLH1	Q9flh1 arabidopsis
847	6	1.4	530	2	O87219	O87219 lactococcus
848	6	1.4	530	5	Q9GPF0	Q9gpf0 dictyosteli
849	6	1.4	531	2	O68307	O68307 nostoc punc
850	6	1.4	531	2	Q44216	Q44216 anabaena sp
851	6	1.4	531	2	Q9HZI0	Q9hzi0 pseudomonas
852	6	1.4	531	2	Q9ZAK2	Q9zak2 anabaena va
853	6	1.4	531	8	O99620	O99620 sarcophyton
854	6	1.4	531	10	O65536	O65536 arabidopsis
855	6	1.4	532	11	O70259	O70259 mus musculu
856	6	1.4	534	8	Q9T2S9	Q9t2s9 solanum tub
857	6	1.4	534	10	Q41445	Q41445 solanum tub
858	6	1.4	534	10	Q9LLE1	Q9lle1 nicotiana t
859	6	1.4	536	8	O20628	O20628 crypthecodi
860	6	1.4	538	1	Q9YCC0	Q9ycc0 aeropyrum p
861	6	1.4	538	2	Q9RC53	Q9rc53 bacillus ha
862	6	1.4	540	2	O69852	O69852 streptomyce
863	6	1.4	540	5	P91411	P91411 caenorhabdi
864	6	1.4	542	3	Q9URQ4	Q9urq4 saccharomyc
865	6	1.4	542	4	Q9NSG4	Q9nsg4 homo sapien
866	6	1.4	542	14	Q9YJ97	Q9yj97 influenza a
867	6	1.4	543	2	Q9PM88	Q9pm88 campylobact
868	6	1.4	543	5	Q19480	Q19480 caenorhabdi
869	6	1.4	545	5	Q9V913	Q9v913 drosophila
870	6	1.4	546	2	Q9I2J7	Q9i2j7 pseudomonas
871	6	1.4	547	5	Q20953	Q20953 caenorhabdi
872	6	1.4	547	5	Q9VWD9	Q9vwd9 drosophila
873	6	1.4	551	5	Q9GRT8	Q9grt8 leishmania
874	6	1.4	552	1	O28554	O28554 archaeoglob
875	6	1.4	555	2	O32382	O32382 streptomyce
876	6	1.4	560	5	Q9VQC0	Q9vqc0 drosophila
877	6	1.4	562	5	O96264	O96264 plasmodium
878	6	1.4	562	10	Q9LRI8	Q9lri8 spinacia ol
879	6	1.4	563	2	O85371	O85371 rhodococcus
880	6	1.4	567	2	Q9K8H0	Q9k8h0 bacillus ha
881	6	1.4	568	5	Q9Y005	Q9y005 priapulus c
882	6	1.4	569	10	O23152	O23152 arabidopsis
883	6	1.4	569	10	Q9LYU8	Q9lyu8 arabidopsis
884	6	1.4	570	5	O97191	O97191 geodia cydo
885	6	1.4	570	10	Q9FK65	Q9fk65 arabidopsis
886	6	1.4	570	14	Q82565	Q82565 influenza a
887	6	1.4	573	4	O94830	O94830 homo sapien
888	6	1.4	573	5	Q9U5N2	Q9u5n2 myzus persi
889	6	1.4	576	5	Q9VXA6	Q9vxa6 drosophila
890	6	1.4	582	3	O74931	O74931 yarrowia li
891	6	1.4	583	11	P70616	P70616 rattus norv
892	6	1.4	586	10	Q9FFD1	Q9ffd1 arabidopsis
893	6	1.4	588	5	Q9VGA8	Q9vga8 drosophila
894	6	1.4	588	10	O22717	O22717 arabidopsis
895	6	1.4	589	2	Q45902	Q45902 coxiella bu
896	6	1.4	590	5	Q9GUG6	Q9gug6 caenorhabdi
897	6	1.4	590	10	O82363	O82363 arabidopsis
898	6	1.4	591	1	O29918	O29918 archaeoglob
899	6	1.4	594	4	Q9UFB9	Q9ufb9 homo sapien
900	6	1.4	595	13	Q9U5F6	Q9u5f6 eptatretus
901	6	1.4	597	2	Q9PAG8	Q9pag8 xylella fas
902	6	1.4	597	14	Q9YUR4	Q9yur4 turkey aden
903	6	1.4	598	3	Q05930	Q05930 saccharomyc
904	6	1.4	601	2	Q9PKY3	Q9pky3 chlamydia m
905	6	1.4	601	5	Q22913	Q22913 caenorhabdi
906	6	1.4	603	8	O79097	O79097 alligator m
907	6	1.4	603	8	Q33790	Q33790 alligator m
908	6	1.4	604	10	Q9FR03	Q9fr03 perilla fru
909	6	1.4	605	5	Q9GYD0	Q9gyd0 leishmania
910	6	1.4	606	10	O22137	O22137 arabidopsis
911	6	1.4	609	3	O94326	O94326 schizosacch
912	6	1.4	611	2	O69506	O69506 mycobacteri
913	6	1.4	612	5	O46102	O46102 drosophila
914	6	1.4	616	2	Q9RYH8	Q9ryh8 deinococcus
915	6	1.4	619	2	Q9F9H4	Q9f9h4 rickettsia
916	6	1.4	621	2	Q9ZFR4	Q9zfr4 burkholderi
917	6	1.4	623	3	O94443	O94443 schizosacch
918	6	1.4	624	2	Q9L4M3	Q9l4m3 prochloroco
919	6	1.4	624	2	Q9K5T7	Q9k5t7 bacillus ha
920	6	1.4	624	2	Q9HYX4	Q9hyx4 pseudomonas
921	6	1.4	632	5	Q9UA36	Q9ua36 caenorhabdi
922	6	1.4	633	10	O22854	O22854 arabidopsis
923	6	1.4	633	10	O65655	O65655 arabidopsis
924	6	1.4	634	2	Q45289	Q45289 brevibacill
925	6	1.4	636	5	Q9Y1J3	Q9y1j3 caenorhabdi
926	6	1.4	636	10	Q9ZSQ0	Q9zsq0 phaseolus a
927	6	1.4	637	5	Q9W0R5	Q9w0r5 drosophila
928	6	1.4	640	2	Q9WZB2	Q9wzb2 thermotoga
929	6	1.4	640	3	O59863	O59863 kluyveromyc
930	6	1.4	641	2	Q9ZA62	Q9za62 neisseria m
931	6	1.4	641	2	Q9KS57	Q9ks57 vibrio chol
932	6	1.4	643	10	Q9LSM5	Q9lsm5 arabidopsis
933	6	1.4	644	8	Q37024	Q37024 hansenula w
934	6	1.4	646	2	Q9PAN5	Q9pan5 xylella fas
935	6	1.4	647	2	Q47911	Q47911 flavobacter
936	6	1.4	649	5	Q22225	Q22225 caenorhabdi
937	6	1.4	650	10	Q9LZJ4	Q9lzj4 arabidopsis
938	6	1.4	652	5	Q07615	Q07615 plasmodium
939	6	1.4	652	5	Q9V411	Q9v411 drosophila
940	6	1.4	653	4	O60271	O60271 homo sapien
941	6	1.4	654	11	Q9Z2G1	Q9z2g1 mus musculu
942	6	1.4	655	5	Q9V9G9	Q9v9g9 drosophila
943	6	1.4	656	10	Q9FEP7	Q9fep7 arabidopsis
944	6	1.4	658	5	Q09946	Q09946 caenorhabdi
945	6	1.4	659	5	Q9N4L1	Q9n4l1 caenorhabdi
946	6	1.4	661	10	Q9SM87	Q9sm87 oryza sativ
947	6	1.4	662	5	Q9VW76	Q9vw76 drosophila
948	6	1.4	662	8	Q9GGV9	Q9ggv9 carex parva
949	6	1.4	663	2	Q9I5D9	Q9i5d9 pseudomonas
950	6	1.4	666	5	O76167	O76167 dugesia jap
951	6	1.4	667	2	Q9X434	Q9x434 streptococc
952	6	1.4	667	2	Q9RPK7	Q9rpk7 streptococc
953	6	1.4	667	11	Q60451	Q60451 cricetulus
954	6	1.4	669	8	Q9GGW1	Q9ggw1 carex nigri
955	6	1.4	669	14	Q9QRV4	Q9qrv4 feline cali
956	6	1.4	669	14	Q9QRV2	Q9qrv2 feline cali
957	6	1.4	669	14	Q9IF03	Q9if03 feline cali
958	6	1.4	670	14	P87674	P87674 feline cali
959	6	1.4	671	2	Q9JXU8	Q9jxu8 neisseria m
960	6	1.4	671	2	Q9JW22	Q9jw22 neisseria m
961	6	1.4	671	14	P87675	P87675 feline cali
962	6	1.4	672	10	Q9FMS5	Q9fms5 arabidopsis
963	6	1.4	673	4	Q9NSG3	Q9nsg3 homo sapien
964	6	1.4	673	14	Q9YX75	Q9yx75 fusarium po
965	6	1.4	674	4	Q9NVZ7	Q9nvz7 homo sapien
966	6	1.4	678	5	O60982	O60982 leishmania
967	6	1.4	678	5	Q9VLG5	Q9vlg5 drosophila
968	6	1.4	679	2	Q9I4A6	Q9i4a6 pseudomonas

969	6	1.4	680	5	Q9VF91	Q9vf91 drosophila
970	6	1.4	681	8	Q9MUI2	Q9mui2 carex scirp
971	6	1.4	682	2	O26013	O26013 helicobacte
972	6	1.4	682	10	Q39948	Q39948 helianthus
973	6	1.4	683	10	O23506	O23506 arabidopsis
974	6	1.4	683	10	Q9LM44	Q9lm44 arabidopsis
975	6	1.4	684	10	O80709	O80709 arabidopsis
976	6	1.4	685	5	Q9N5E3	Q9n5e3 caenorhabdi
977	6	1.4	687	5	Q20900	Q20900 caenorhabdi
978	6	1.4	687	8	Q9GGW6	Q9ggw6 carex antho
979	6	1.4	688	8	Q9GGW4	Q9ggw4 carex elyno
980	6	1.4	690	5	Q27598	Q27598 drosophila
981	6	1.4	690	5	Q9V896	Q9v896 drosophila
982	6	1.4	691	1	Q9Y9K9	Q9y9k9 aeropyrum p
983	6	1.4	691	11	Q9R031	Q9r031 mus musculu
984	6	1.4	695	8	Q9MUI4	Q9mui4 carex echin
985	6	1.4	695	8	Q9GGV1	Q9ggv1 kobresia ro
986	6	1.4	695	11	Q9Z0U0	Q9z0u0 mus musculu
987	6	1.4	695	11	Q9R036	Q9r036 mus musculu
988	6	1.4	695	11	Q9QZ72	Q9qz72 mus musculu
989	6	1.4	696	4	O95719	O95719 homo sapien
990	6	1.4	696	4	Q9UBH6	Q9ubh6 homo sapien
991	6	1.4	696	6	Q9TU72	Q9tu72 mustela vis
992	6	1.4	696	8	Q9GGV4	Q9ggv4 kobresia fr
993	6	1.4	696	8	Q9GGU1	Q9ggu1 eriophorum
994	6	1.4	696	11	Q9R035	Q9r035 mus dunni.
995	6	1.4	696	11	Q9R034	Q9r034 mus musculu
996	6	1.4	696	11	Q9R033	Q9r033 cricetulus
997	6	1.4	696	11	Q9R032	Q9r032 mus spretus
998	6	1.4	696	11	Q9QZ71	Q9qz71 mus dunni.
999	6	1.4	696	11	Q9QZ70	Q9qz70 cricetulus
1000	6	1.4	698	8	Q9GGW0	Q9ggw0 carex obtus

ALIGNMENTS

```
RESULT   1
Q9VSE7
ID   Q9VSE7      PRELIMINARY;      PRT;   426 AA.
AC   Q9VSE7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CG7476 PROTEIN.
GN   CG7476 OR CG7674.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003556; AAF50475.1; -.
DR   FlyBase; FBgn0035847; CG7476.
DR   InterPro; IPR000832; -.
SQ   SEQUENCE   426 AA;  49630 MW;  BF11D7EBE82F0561 CRC64;
```

```
  Query Match           100.0%;  Score 426;  DB 5;  Length 426;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MRLPWVIFCTVLLLIFTNNSNADIPGCNYYDTVDISYIERQNDSYLYDDIEIPASLTGYY 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MRLPWVIFCTVLLLIFTNNSNADIPGCNYYDTVDISYIERQNDSYLYDDIEIPASLTGYY 60

Qy      61 EFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLKEELARFKPYIYF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 EFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLKEELARFKPYIYF 120

Qy     121 TYMDLQARVPLTDMAIIRDEFFDCDEMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TYMDLQARVPLTDMAIIRDEFFDCDEMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFW 180

Qy     181 VTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVY 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 VTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLEILIITMICFVLTIAVY 240

Qy     241 LYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLSV 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 LYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLSV 300

Qy     301 ISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWENDPSKWNWLPLVG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMTGSIYIVNQIWENDPSKWNWLPLVG 360

Qy     361 FIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGINKFTNEEEGRINCI 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 FIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGINKFTNEEEGRINCI 420

Qy     421 NFDSQT 426
           ||||||
Db     421 NFDSQT 426
```

```
RESULT   2
Q69093
ID   Q69093      PRELIMINARY;      PRT;   179 AA.
AC   Q69093;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL PROTEIN (FRAGMENT).
OS   Herpes simplex virus (type 2).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
```

```
OC   Alphaherpesvirinae; Simplexvirus.
OX   NCBI_TaxID=10310;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HG52;
RX   MEDLINE=87071654; PubMed=3024391;
RA   Bzik D.J., Debroy C., Fox B.A., Pederson N.E., Person S.;
RT   "The nucleotide sequence of the gB glycoprotein gene of HSV-2 and
RT   comparison with the corresponding gene of HSV-1.";
RL   Virology 155:322-333(1986).
DR   EMBL; M14923; AAA66439.1; -.
DR   InterPro; IPR000501; -.
DR   Pfam; PF01366; PRTP; 1.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   179 AA;  19633 MW;  BB25DED49FBE023C CRC64;


  Query Match             1.9%;  Score 8;  DB 14;  Length 179;
  Best Local Similarity   100.0%;  Pred. No. 6.3;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db      34 LKEELARF 41


RESULT   3
Q04459
ID   Q04459      PRELIMINARY;      PRT;   266 AA.
AC   Q04459;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PROCESSING AND TRANSPORT PROTEIN (INFECTED CELL PROTEIN 18.5)
DE   (ICP18.5 PROTEIN) (FRAGMENT).
GN   UL28.
OS   Herpesvirus ateles (type 1 / strain Lennette).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae.
OX   NCBI_TaxID=35243;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93228440; PubMed=8385913;
RA   Eberle R., Black D.;
RT   "Sequence analysis of herpes simplex virus gB gene homologs of two
RT   platyrrhine monkey alpha-herpesviruses.";
RL   Arch. Virol. 129:167-182(1993).
CC   -!- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLOCATION OF THE VIRUS
CC       GLYCOPROTEINS TO MEMBRANES.
CC   -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL28,
CC       EHV-1 32, PRV ICP18.5, EBV BALF3, AND VZV 30.
DR   EMBL; M95785; AAA43838.1; -.
DR   InterPro; IPR000501; -.
DR   Pfam; PF01366; PRTP; 1.
KW   Late protein.
FT   NON_TER       1      1
SQ   SEQUENCE   266 AA;  29883 MW;  22F5F52FC1474B3E CRC64;


  Query Match             1.9%;  Score 8;  DB 14;  Length 266;
  Best Local Similarity   100.0%;  Pred. No. 8.8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     116 LKEELARF 123


RESULT   4
O94006
ID   O94006      PRELIMINARY;    PRT;   270 AA.
AC   O94006;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   POTENTIAL MEMBRANE PROTEIN.
GN   CA20C1.19.
OS   Candida albicans (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5476;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1161;
RA   Oliver K., Harris D.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1161;
RA   Barrell B.G., Rajandream M.A.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1161;
RX   MEDLINE=97435544; PubMed=9290243;
RA   Tait E., Simon M.C., King S., Brown A.J., Gow N.A.R., Shaw D.J.;
RT   "A Candida albicans genome project: cosmid contigs, physical mapping,
RT   and gene isolation.";
RL   Fungal Genet. Biol. 21:308-314(1997).
DR   EMBL; AL033391; CAA21940.1; -.
SQ   SEQUENCE   270 AA;  29915 MW;  46940D75D487460D CRC64;


  Query Match             1.9%;  Score 8;  DB 3;  Length 270;
  Best Local Similarity   100.0%;  Pred. No. 9;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 TVLLLIFT 17
           ||||||||
Db      16 TVLLLIFT 23


RESULT   5
Q69261
ID   Q69261         PRELIMINARY;    PRT;   362 AA.
AC   Q69261;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ANALOGOUS ORF TO HSV1 ICP18.5. PRECURSOR (FRAGMENT).
OS   Equine herpesvirus 1.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=10326;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89279217; PubMed=2543744;
RA   Whalley J.M., Robertson G.R., Scott N.A., Hudson G.C., Bell C.W.,
RA   Woodworth L.M.;
RT   "Identification and nucleotide sequence of a gene in equine
RT   herpesvirus 1 analogous to the herpes simplex virus gene encoding the
RT   major envelope glycoprotein gB.";
RL   J. Gen. Virol. 70:383-394(1989).
DR   EMBL; D00401; BAA00303.1; -.
DR   InterPro; IPR000501; -.
DR   Pfam; PF01366; PRTP; 1.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL     <317    362       POTENTIAL.
SQ   SEQUENCE   362 AA;  39708 MW;  7875DBED279AA2C1 CRC64;


  Query Match             1.9%;  Score 8;  DB 14;  Length 362;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy     107 LKEELARF 114
           ||||||||
Db     225 LKEELARF 232


RESULT   6
Q9VRN2
ID   Q9VRN2      PRELIMINARY;      PRT;   482 AA.
AC   Q9VRN2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CG17795 PROTEIN.
GN   CG17795.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003565; AAF50761.1; -.
DR   FlyBase; FBgn0035623; CG17795.
DR   InterPro; IPR000276; -.
DR   InterPro; IPR000832; -.
DR   Pfam; PF00002; 7tm_2; 1.
SQ   SEQUENCE   482 AA;  56250 MW;  E14CE324A7E09163 CRC64;


  Query Match             1.9%;  Score 8;  DB 5;  Length 482;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     234 VLTIAVYL 241
           ||||||||
Db     226 VLTIAVYL 233


RESULT   7
Q9W0R4
ID   Q9W0R4      PRELIMINARY;      PRT;   513 AA.
AC   Q9W0R4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   MTH PROTEIN.
GN   MTH OR CG6936.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003468; AAF47379.1; -.
DR   FlyBase; FBgn0023000; mth.
DR   InterPro; IPR000832; -.
DR   Pfam; PF00002; 7tm_2; 1.
SQ   SEQUENCE   513 AA;  59497 MW;  66C8511E97E5E975 CRC64;


  Query Match             1.9%;  Score 8;  DB 5;  Length 513;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy     234 VLTIAVYL 241
           ||||||||
Db     230 VLTIAVYL 237


RESULT   8
O97148
ID   O97148      PRELIMINARY;      PRT;   514 AA.
AC   O97148;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PUTATIVE G-PROTEIN-COUPLED RECEPTOR.
GN   MTH OR CG6936.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON S;
RX   MEDLINE=99011353; PubMed=9794765;
RA   Lin Y.J., Seroude L., Benzer S.;
RT   "Extended life-span and stress resistance in the Drosophila mutant
RT   methuselah.";
RL   Science 282:943-946(1998).
DR   EMBL; AF109308; AAD16981.1; -.
DR   FlyBase; FBgn0023000; mth.
DR   InterPro; IPR000832; -.
DR   Pfam; PF00002; 7tm_2; 1.
KW   Receptor.
SQ   SEQUENCE   514 AA;  59644 MW;  F2A5370CECCD1D5C CRC64;


  Query Match             1.9%;  Score 8;  DB 5;  Length 514;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     234 VLTIAVYL 241
           ||||||||
Db     230 VLTIAVYL 237


RESULT   9
Q9GT56
ID   Q9GT56      PRELIMINARY;      PRT;   517 AA.
AC   Q9GT56;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BG1013;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280583; AAG21991.1; -.
SQ   SEQUENCE   517 AA;  59954 MW;  839CBF03F93283CF CRC64;


  Query Match             1.9%;  Score 8;  DB 5;  Length 517;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     29 YYDTVDIS 36
          ||||||||
Db     34 YYDTVDIS 41


RESULT  10
Q9GT55
ID   Q9GT55      PRELIMINARY;      PRT;   517 AA.
AC   Q9GT55;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YAK_2;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280584; AAG21992.1; -.
SQ   SEQUENCE   517 AA;  60094 MW;  D7D724AB77322728 CRC64;


  Query Match             1.9%;  Score 8;  DB 5;  Length 517;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     29 YYDTVDIS 36
          ||||||||
Db     34 YYDTVDIS 41


RESULT  11
Q9GT53
ID   Q9GT53      PRELIMINARY;      PRT;   517 AA.
AC   Q9GT53;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YAK_30;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280586; AAG21994.1; -.
SQ   SEQUENCE   517 AA;  59945 MW;  5880753D7E9EE545 CRC64;


  Query Match             1.9%;  Score 8;  DB 5;  Length 517;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     29 YYDTVDIS 36
          ||||||||
Db     34 YYDTVDIS 41
```

```
RESULT  12
Q9GT52
ID   Q9GT52      PRELIMINARY;      PRT;   517 AA.
AC   Q9GT52;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YAK_4;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280587; AAG21995.1; -.
SQ   SEQUENCE   517 AA;  59938 MW;  872FC1ACFDE837B6 CRC64;

  Query Match             1.9%;  Score 8;  DB 5;  Length 517;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      29 YYDTVDIS 36
           ||||||||
Db      34 YYDTVDIS 41


RESULT  13
Q9GT51
ID   Q9GT51      PRELIMINARY;      PRT;   517 AA.
AC   Q9GT51;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   METHUSELAH.
GN   MTH.
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YAK_5;
RX   MEDLINE=20461461; PubMed=10995474;
RA   Schmidt P.S., Duvernell D.D., Eanes W.F.;
RT   "Adaptive evolution of a candidate gene for aging in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:10861-10865(2000).
DR   EMBL; AF280588; AAG21996.1; -.
SQ   SEQUENCE   517 AA;  60032 MW;  7D9726B5D4445981 CRC64;

  Query Match             1.9%;  Score 8;  DB 5;  Length 517;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      29 YYDTVDIS 36
           ||||||||
Db      34 YYDTVDIS 41


RESULT  14
O39274
ID   O39274      PRELIMINARY;      PRT;   773 AA.
AC   O39274;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   COUNTERPART OF HSV-1 GENE UL28 AND VZV GENE 30.
GN   32.
OS   Equine herpesvirus 4.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicellovirus.
OX   NCBI_TaxID=10331;
RN   [1]
RP   SEQUENCE OF 639-773 FROM N.A.
RC   STRAIN=NS80567;
RX   MEDLINE=89125704; PubMed=2915378;
RA   Riggio M.P., Cullinane A.A., Onions D.E.;
RT   "Identification and nucleotide sequence of the glycoprotein gB gene of
RT   equine herpesvirus 4.";
RL   J. Virol. 63:1123-1133(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NS80567;
RX   MEDLINE=98264497; PubMed=9603335;
RA   Telford E.A.R., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
RT   "The DNA sequence of equine herpesvirus-4.";
RL   J. Gen. Virol. 79:1197-1203(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NS80567;
RA   Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF030027; AAC59548.1; -.
DR   InterPro; IPR000501; -.
DR   Pfam; PF01366; PRTP; 1.
SQ   SEQUENCE   773 AA;  85317 MW;  A779945277E0D956 CRC64;

  Query Match             1.9%;  Score 8;  DB 14;  Length 773;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     636 LKEELARF 643


RESULT  15
P89451
ID   P89451      PRELIMINARY;      PRT;   785 AA.
AC   P89451;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   HERPES SIMPLEX VIRUS TYPE 2 (STRAIN HG52), COMPLETE GENOME.
GN   UL28.
OS   Herpes simplex virus (type 2).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
OX   NCBI_TaxID=10310;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HG52;
RX   MEDLINE=87111457; PubMed=3027242;
RA   McGeoch D.J., Moss H.W., McNab D., Frame M.C.;
RT   "DNA sequence and genetic content of the HindIII l region in the short
RT   unique component of the herpes simplex virus type 2 genome:
RT   identification of the gene encoding glycoprotein G, and evolutionary
RT   comparisons.";
RL   J. Gen. Virol. 68:19-38(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HG52;
RX   MEDLINE=90278430; PubMed=2161906;
RA   Everett R., Fenwick M.;
RT   "Comparative DNA sequence analysis of the host shutoff genes of
RT   different strains of herpes simplex virus: type 2 strain HG52 encodes
```

```
RT   a truncated UL41 product.";
RL   J. Gen. Virol. 71:1387-1390(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HG52;
RX   MEDLINE=92113549; PubMed=1662697;
RA   McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;
RT   "Comparative sequence analysis of the long repeat regions and
RT   adjoining parts of the long unique regions in the genomes of herpes
RT   simplex viruses types 1 and 2.";
RL   J. Gen. Virol. 72:3057-3075(1991).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HG52;
RX   MEDLINE=92356101; PubMed=1322965;
RA   Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;
RT   "A novel herpes simplex virus gene (UL49A) encodes a putative membrane
RT   protein with counterparts in other herpesviruses.";
RL   J. Gen. Virol. 73:2167-2171(1992).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HG52;
RA   Dolan A.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z86099; CAB06753.1; -.
DR   InterPro; IPR000501; -.
DR   Pfam; PF01366; PRTP; 1.
SQ   SEQUENCE   785 AA;  85240 MW;  246988E41997DF62 CRC64;


  Query Match             1.9%;  Score 8;  DB 14;  Length 785;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     107 LKEELARF 114
           ||||||||
Db     640 LKEELARF 647


Search completed: October 30, 2001, 15:33:35
Job time: 221 sec
```

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2001, 10:29:58 ; Search time 2043.44 Seconds
(without alignments)
11626.699 Million cell updates/sec

Title: US-09-619-049-854
Perfect score: 1536
Sequence: 1 atgaggcttccttgggtgat.....ggactctaatgatggattct 1536

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*
44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vi1:*
59: gb_vi2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_ro1:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1206	78.5	83545	65	AC020432	AC020432 Drosophil
c	2	1206	78.5	281146	4	AE003556	AE003556 Drosophil
	3	846.6	55.1	46913	68	AC023733	AC023733 Drosophil
	4	460	29.9	4616	5	AF053725	AF053725 Drosophil
c	5	436	28.4	164659	61	AC010037	AC010037 Drosophil
	6	196	12.8	42611	64	AC017388	AC017388 Drosophil
	7	196	12.8	273744	4	AE003558	AE003558 Drosophil
	8	167.6	10.9	25416	63	AC013100	AC013100 Drosophil

9	167.6	10.9	188967	4	AC007754	AC007754 Drosophil
10	167.6	10.9	225668	5	AE003699	AE003699 Drosophil
11	144.8	9.4	153511	61	AC010557	AC010557 Drosophil
12	129.4	8.4	121857	30	AC008360	Ac008360 Drosophil
13	124.8	8.1	1948	5	AF109308	AF109308 Drosophil
14	94.8	6.2	143393	63	AC015222	AC015222 Drosophil
15	94.8	6.2	182494	4	AC009343	AC009343 Drosophil
16	94.8	6.2	259934	5	AE003828	AE003828 Drosophil
17	92	6.0	1928	5	AF280584	AF280584 Drosophil
18	91.6	6.0	1928	5	AF280587	AF280587 Drosophil
19	90.8	5.9	1928	5	AF280586	AF280586 Drosophil
20	90.4	5.9	1928	5	AF280583	AF280583 Drosophil
21	89.6	5.8	1928	5	AF280589	AF280589 Drosophil
22	89.2	5.8	1928	5	AF280585	AF280585 Drosophil
23	88.8	5.8	1928	5	AF280588	AF280588 Drosophil
24	86.2	5.6	196598	60	AC008004	AC008004 Drosophil
25	84.4	5.5	1927	5	AF280569	AF280569 Drosophil
26	84.4	5.5	1927	5	AF280574	AF280574 Drosophil
27	84.4	5.5	1927	5	AF280576	AF280576 Drosophil
28	84.4	5.5	1927	5	AF280578	AF280578 Drosophil
29	84.4	5.5	1927	5	AF280579	AF280579 Drosophil
30	84.4	5.5	1927	5	AF280580	AF280580 Drosophil
31	82.8	5.4	1917	5	AF280581	AF280581 Drosophil
32	82.8	5.4	1927	5	AF280552	AF280552 Drosophil
33	82.8	5.4	1927	5	AF280553	AF280553 Drosophil
34	82.8	5.4	1927	5	AF280554	AF280554 Drosophil
35	82.8	5.4	1927	5	AF280555	AF280555 Drosophil
36	82.8	5.4	1927	5	AF280556	AF280556 Drosophil
37	82.8	5.4	1927	5	AF280557	AF280557 Drosophil
38	82.8	5.4	1927	5	AF280558	AF280558 Drosophil
39	82.8	5.4	1927	5	AF280559	AF280559 Drosophil
40	82.8	5.4	1927	5	AF280560	AF280560 Drosophil
41	82.8	5.4	1927	5	AF280561	AF280561 Drosophil
42	82.8	5.4	1927	5	AF280562	AF280562 Drosophil
43	82.8	5.4	1927	5	AF280563	AF280563 Drosophil
44	82.8	5.4	1927	5	AF280564	AF280564 Drosophil
45	82.8	5.4	1927	5	AF280565	AF280565 Drosophil

ALIGNMENTS

```
RESULT   1
AC020432
LOCUS       AC020432    83545 bp    DNA             HTG       03-JAN-2000
DEFINITION  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
            pieces.
ACCESSION   AC020432
VERSION     AC020432.1  GI:6664465
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 83545)
  AUTHORS   Adams,M. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     This sequence was identified as CDM:10213149 by the submitter.
            For more information on this record e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .83545
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
BASE COUNT    24666 a  18278 c  17979 g  22622 t
ORIGIN
```

```
  Query Match             78.5%;  Score 1206;  DB 65;  Length 83545;
  Best Local Similarity   83.7%;  Pred. No. 0;
  Matches 1536;  Conservative    0;  Mismatches    0;  Indels  300;  Gaps    3;

Qy       1 atgaggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   33705 ATGAGGCTTCCTTGGGTGATTTTCTGTACAGTTCTGCTTTTAATTTTTACGAACAACTCA 33764

Qy      61 aatgcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaagg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   33765 AATGCAGATATTCCCGGGTGCAACTACTACGACACGGTTGATATCTCATACATTGAAAGG 33824

Qy     121 caaaacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactac 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   33825 CAAAACGATTCGTATTTATATGATGACATCGAAATTCCTGCTAGCTTAACTGGATACTAC 33884

Qy     181 gagttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgt 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   33885 GAGTTCAGGCAGTTTGGGGACGGTTCGATTACGCCGATTGAAAAGCATTTAAGGGCCTGT 33944

Qy     241 gtttgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttttggccaac 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   33945 GTTTGCAGCGTGCGGCCCTGTATTCGAATCTGCTGCCCAGCCAAGAACTTTTTGGCCAAC 34004

Qy     301 ggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttc 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34005 GGAAAATGCGATGATGGTCTCAAAGAGGAGCTCGCCCGGTTCAAACCCTATATATACTTC 34064

Qy     361 acatacatggacctacaggcacgagtaccactcaccgatatggctattatcagagatgag 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34065 ACATACATGGACCTACAGGCACGAGTACCACTCACCGATATGGCTATTATCAGAGATGAG 34124

Qy     421 ttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggtt 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34125 TTCTTTGATTGTGATGAAATGATTTACATCAGCGACTTTAACTATTTCTTAGAAGAGGTT 34184

Qy     481 agcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaatttttgg 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34185 AGCATTCAAATCTTCAATAAATGTGGACTTATAGTTTGGTTCCAGGATGGAAAATTTTGG 34244

Qy     541 gttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaacctttgat 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34245 GTTACCGTTGACCTCTTCATGGAAAAACAGGACTACTGCTTGTACCGGCACAACTTTGAT 34304

Qy     601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34305 TCGGATTTTCCAAAGTCCATGTGGATAATACGACACCGCTGCACAAGCCACATATCTCCT 34364

Qy     661 ggatccttagaga---------------------------------------------- 673
           ||||||||||||
Db   34365 GGATCCTTAGAGAGTAATTCAATTATTATTTATAACAAAAGTCATCAATATCGTTTTCAT 34424

Qy     674 ---------------------------------------------------------- 673

Db   34425 AGCTATATAGAAAGAAACGAAAATAACTCCTTATTGGTGTGCAAACTGCTGACTGCATTC 34484

Qy     674 ---------------------------------------------------------- 673

Db   34485 TGTGTGCCCTTTGCAAGTTTTTTAATTACAATAAGCTGGTTTTTTATTTCTTATATTTTC 34544

Qy     674 --------ttctaattataacaatgatatgctttgtcctaacaatcgcagtatatctata 725
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34545 TTTTTCAGTTCTAATTATAACAATGATATGCTTTGTCCTAACAATCGCAGTATATCTATA 34604

Qy     726 cattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcgaggtt 785
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34605 CATTAAGAAGCTGCGAAATGTTACTGGCAAGTGCATTGTATGCTGTATAGTTTCGAGGTT 34664

Qy     786 tatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgctctcc 845
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34665 TATCCAGTGCTTGATCATGATACTAGATCATTTAAATCTATTGAATGGCATTTGCTCTCC 34724
```

```
Qy     846 agct---------------------------------------------------------- 849
           ||||
Db   34725 AGCTGGTTAGTTCGAATTGGAGATATCTCTGGAAGCTTTAAATGTTATATCGTTGACTTC 34784

Qy     850 -----ggttacagctcgcacttttttccggatggcttccaacctctggctctccgtcatca 904
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34785 TTGCAGGTTACAGCTCGCACTTTTTCCGGATGGCTTCCAACCTCTGGCTCTCCGTCATCA 34844

Qy     905 gctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaactatcggttcc 964
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34845 GCTACCATACGTGGAAAGTCTTGACGTCGCTCAATCGAGTCGACCCTAACTATCGGTTCC 34904

Qy     965 tgcggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttatatag 1024
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34905 TGCGGTACAACGCCTTCGTCTGGAGCACAGCCGCAATCATGACGGGAAGTATTTATATAG 34964

Qy    1025 ttaatcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggtttta 1084
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   34965 TTAATCAGATTTGGGAAAACGATCCCAGTAAATGGAACTGGTTGCCTCTGGTCGGTTTTA 35024

Qy    1085 ttcggtgctcggtcaa-------------------------------------------- 1100
           ||||||||||||||||
Db   35025 TTCGGTGCTCGGTCAAAGGTAAACTTTCAATGCGATTGCTTTTAAATCGTTACCATGGTT 35084

Qy    1101 --------------------agattggcacccatccgtctggatctatataagtggaccg 1140
                               ||||||||||||||||||||||||||||||||||||||||
Db   35085 ACGTATTGCTTTTCCATATCAGATTGGCACCCATCCGTCTGGATCTATATAAGTGGACCG 35144

Qy    1141 tcgctggccctgagcactttcaatgtcgccatgttcgccctgacagccatttacattagg 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35145 TCGCTGGCCCTGAGCACTTTCAATGTCGCCATGTTCGCCCTGACAGCCATTTACATTAGG 35204

Qy    1201 aaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggattaactgcata 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35205 AAAGTGAAGGGGGGTATAAATAAGTTCACAAATGAGGAGGAGGGAAGGATTAACTGCATA 35264

Qy    1261 aactttgacagccagacgtaagatatgataaaaccttggaccttggtaaactagttatcc 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35265 AACTTTGACAGCCAGACGTAAGATATGATAAAACCTTGGACCTTGGTAAACTAGTTATCC 35324

Qy    1321 ttgccgtattatccacatagttacctacagttcctgcggctctccatcgtgatgggcctt 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35325 TTGCCGTATTATCCACATAGTTACCTACAGTTCCTGCGGCTCTCCATCGTGATGGGCCTT 35384

Qy    1381 acttggatattcaatgtcattccgtattctgcacggctccacattttctgggaatgggtc 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35385 ACTTGGATATTCAATGTCATTCCGTATTCTGCACGGCTCCACATTTTCTGGGAATGGGTC 35444

Qy    1441 gggataatatccgagtattttcacagcgcgtttggaattgttttgttcgttctgctcgtc 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35445 GGGATAATATCCGAGTATTTTCACAGCGCGTTTGGAATTGTTTTGTTCGTTCTGCTCGTC 35504

Qy    1501 ctgaagcgcagcacatggactctaatgatggattct 1536
           |||||||||||||||||||||||||||||||||||
Db   35505 CTGAAGCGCAGCACATGGACTCTAATGATGGATTCT 35540


RESULT   2
AE003556/c
LOCUS       AE003556   281146 bp    DNA             INV       06-OCT-2000
DEFINITION  Drosophila melanogaster genomic scaffold 142000013386050 section 12
            of 54, complete sequence.
ACCESSION   AE003556 AE002602
VERSION     AE003556.2  GI:10728058
KEYWORDS    HTG.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 281146)
  AUTHORS   Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
            Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
            George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
            Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
            Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
            Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
            Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
            Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
            Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
            Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
            Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
            Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
            Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
            Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
            Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
            Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S.,
            Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S.,
            Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
            Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
            Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
            Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
            Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
            Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
            Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
            McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
            Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
            Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
            Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
            Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
            Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
            Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
            Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,
            Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
            Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,
            Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
            Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
            Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
            Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
  TITLE     The genome sequence of Drosophila melanogaster
  JOURNAL   Science 287 (5461), 2185-2195 (2000)
  MEDLINE   20196006
REFERENCE   2  (bases 1 to 281146)
  AUTHORS   Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     On Oct 9, 2000 this sequence version replaced gi:7295117.
FEATURES             Location/Qualifiers
     source          1. .281146
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="3L"
     mRNA            complement(join(2028. .3137,3574. .3871,5578. .5744,
                     6159. .6388))
                     /gene="bip1"
                     /note="Nucleotide sequence of the Celera sequence differs
                     from the published sequence for this transcript."
                     /product="CT23173"
                     /db_xref="FLYBASE:FBan0007574"
                     /db_xref="FLYBASE:FBgn0026263"
     gene            complement(<2028. .>6388)
                     /gene="bip1"
                     /note="CG7574"
                     /map="66B4-66B5"
                     /db_xref="FLYBASE:FBan0007574"
                     /db_xref="FLYBASE:FBgn0026263"
     CDS             complement(join(2343. .3137,3574. .3871,5578. .5744))
                     /gene="bip1"
                     /note="bip1 gene product; Nucleotide sequence of the
                     Celera sequence differs from the published sequence for
                     this transcript"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0007574"
                     /db_xref="FLYBASE:FBgn0026263"
                     /protein_id="AAF50496.1"
```

```
                     /db_xref="GI:7295171"
                     /translation="MTNTITTLRICQAVSLNPSANGPEPVLPPSSRICLMGCQDVKPN
                     DLRRFPVHDTDRCNHWLRHFGVSYKLVDQLGGLQKLRICSRHFSQRQEMAKTRIKMTP
                     LRKSSSSKGIVLPINAAGGSVIAGALAGGGAGTGSGAVKSAVASSTTIELVESPNRH
                     QAKKPNTGIMPMPQRSHSISSSASSDVLSINSDYEASLPLSALEDSQNGLNRLENGQA
                     KKRKLYVITEKSDETKGQTLETNGQPKKRKMFVVVEQDKDVAGKKLMIMADKPQTNLT
                     QHLEKLLPELLNRIQDKEQKQIKPAPVVTKAKPVVNNPRPHITLPPKKNPTNFAKITS
                     TPSSVPSSIQMPLPIQNTKILKSENSELTMNSDEIETNINLSPDFRFLMKILPKLEQL
                     PEPHKQNVKRSIQIFVEKSYSIYGNKD"
mRNA                 <19498. .>20010
                     /gene="CG13681"
                     /product="CT33117"
                     /db_xref="FLYBASE:FBan0013681"
                     /db_xref="FLYBASE:FBgn0035832"
                     /evidence=not_experimental
gene                 <19498. .>20010
                     /gene="CG13681"
                     /map="66B5-66B6"
                     /db_xref="FLYBASE:FBan0013681"
                     /db_xref="FLYBASE:FBgn0035832"
                     /evidence=not_experimental
CDS                  19498. .20010
                     /gene="CG13681"
                     /note="CG13681 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0013681"
                     /db_xref="FLYBASE:FBgn0035832"
                     /evidence=not_experimental
                     /protein_id="AAF50495.1"
                     /db_xref="GI:7295170"
                     /translation="MNSPQLAPTTAQSHHAESADKTAPPGNKYAPGRRLAARQTLRLA
                     IPKAQPGDPLNHPTSVGLFRPSPSQLPPAPILKRVPAVSGGCSGIPKSASPLESHHLP
                     NVTPTNAHTRLGYGMQGAGGGSSCSSASSASNFLGVRKPSLTLNSSDIRNNNDLFAFP
                     KDMSPVMSGK"
mRNA                 complement(join(35698. .35900,35969. .36136,36196. .36712,
                     36772. .36962,37025. .37133,37194. .37325,37382. .38665,
                     38732. .>39472))
                     /gene="CG7565"
                     /product="CT23145"
                     /db_xref="FLYBASE:FBan0007565"
                     /db_xref="FLYBASE:FBgn0035833"
gene                 complement(<35698. .>39472)
                     /gene="CG7565"
                     /map="66B6-66B7"
                     /db_xref="FLYBASE:FBan0007565"
                     /db_xref="FLYBASE:FBgn0035833"
CDS                  complement(join(35698. .35900,35969. .36136,36196. .36712,
                     36772. .36962,37025. .37133,37194. .37325,37382. .38665,
                     38732. .39337))
                     /gene="CG7565"
                     /note="CG7565 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0007565"
                     /db_xref="FLYBASE:FBgn0035833"
                     /protein_id="AAF50494.1"
                     /db_xref="GI:7295169"
                     /translation="MVQVCKRICNLLLLATAMSSAYADVTTQNALLVGSKKHKETSPD
                     NSVGGSISPNLVCHKMLRHVFENATPRDEQQAGVFEEYKPPPDAVEPLEEEAYLWNCL
                     QACCEKPRNGSSACNVVLVFKAKCYHIRCQSNEACLPKLRVRMPNEKVQMVLVNPLGD
                     ATWPQLLKAEAAKQNAEILPYDEAALNFWKQPRRLSYLARNQETPVYEDEDFPLADKR
                     MNQMIFQPDENDVLANEELGYYDSNAKFTTCDMETPCPPPQQCVPLQPNAVRGVCTCP
                     EGFVWNKQRKCVMAAVPYSSYLTSNEAGQQEAAASENSPEVSTPPLKAEQNKDIVVSV
                     MSKEVRLPEQEVTLAAFTVPDEQTSDTKYKYLWTLISQPKGPMNGTISDQSKSKVKLS
                     NLSEGLYTFKVTVTGDNGTFGEATANVTVLPENRINQPPQVIISPREQIIRQPTTNAI
                     LDGSTSTDDDKITNWHWEVISGPIGYQPVLPEVNTLQLDLTSPGNYTFKLTVTDSNNV
                     TNSTTATIAVLKETDYAPVANAGDAVILYLPNNNVTLNGTASSDDHEIVAWEWTKDAS
                     DEAKAVDMQNTRTPYVQLSNLEEGMYTFVLKVTDGSGQSSTAKVHVFVKPPTNSPPVA
                     EAGSNTTTSLPINWVLLNGSDSKDDIGIKSYLWKQLSGPNNAVILKSNSSIANATSLT
                     LGLYEFELTVADENNNTATDTTWVKIVQERNAAPIANAGGDHTVTLPATAIYFNGSKS
                     WDDLAVVKYLWTRDEHSLAAGVIVADTDKEPVMILTNLVQQRYVFTLTVSDDQGLTSS
                     DTVSVNVRRDPKLLNLVQMTLPMGISVLVQSELDSVVQKLQLLLGDENKIQVRELKYD
                     LHTDATVLVFYVNDGQGKALDGLQVERQLRTQLQKDASILGAFAVDIRTSVCQSDCSG
                     HGSCNPITRACICEAFWMPSAGYFFNNQEANCDWSILYVFVGVIVGCLLLSGVFWGIA
                     CACRQSKKPRLRQKVQKYSLIGNKDEEAANYSRNTSLTESETDSDVLFETRTKSNGLG
                     KHKSHNSHSHGHGSSGGSGSSGRDGHKYGTQKLGRKIKA"
mRNA                 join(<40759. .41037,41176. .>41187)
                     /gene="CG13680"
                     /product="CT33116"
                     /db_xref="FLYBASE:FBan0013680"
                     /db_xref="FLYBASE:FBgn0040834"
                     /evidence=not_experimental
gene                 <40759. .>41187
                     /gene="CG13680"
                     /map="66B6-66B7"
                     /db_xref="FLYBASE:FBan0013680"
                     /db_xref="FLYBASE:FBgn0040834"
                     /evidence=not_experimental
CDS                  join(40759. .41037,41176. .41187)
                     /gene="CG13680"
                     /note="CG13680 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0013680"
                     /db_xref="FLYBASE:FBgn0040834"
                     /evidence=not_experimental
                     /protein_id="AAF50493.1"
                     /db_xref="GI:7295168"
                     /translation="MTDFSDSSPSGVLSPPTSLIENFHKKNSSPQPLGSSSAISTLTP
                     TTLNGSVVGGGVGGAGHHSIENGRQINTSSSSCSSLKETSHQSTSTSQQNTD"
mRNA                 join(44247. .44576,44671. .44872,46046. .46425,46846. .47236,
                     47295. .>47339)
                     /gene="CG7905"
                     /product="CT23884"
                     /db_xref="FLYBASE:FBan0007905"
                     /db_xref="FLYBASE:FBgn0035834"
gene                 <44247. .>47339
                     /gene="CG7905"
                     /map="66B6-66B7"
                     /db_xref="FLYBASE:FBan0007905"
                     /db_xref="FLYBASE:FBgn0035834"
CDS                  join(46169. .46425,46846. .47236,47295. .47339)
                     /gene="CG7905"
                     /note="CG7905 gene product"
                     /codon_start=1
```

```
Query Match             78.5%;  Score 1206;  DB 4;  Length 281146;
Best Local Similarity   83.7%;  Pred. No. 0;
Matches 1536;  Conservative    0;  Mismatches    0;  Indels  300;  Gaps    3;

Qy       1 atgaggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  128217 ATGAGGCTTCCTTGGGTGATTTTCTGTACAGTTCTGCTTTTAATTTTTACGAACAACTCA 128158

Qy      61 aatgcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaagg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  128157 AATGCAGATATTCCCGGGTGCAACTACTACGACACGGTTGATATCTCATACATTGAAAGG 128098

Qy     121 caaaacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactac 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  128097 CAAAACGATTCGTATTTATATGATGACATCGAAATTCCTGCTAGCTTAACTGGATACTAC 128038

Qy     181 gagttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgt 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  128037 GAGTTCAGGCAGTTTGGGGACGGTTCGATTACGCCGATTGAAAAGCATTTAAGGGCCTGT 127978

Qy     241 gtttgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttggccaac 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  127977 GTTTGCAGCGTGCGGCCCTGTATTCGAATCTGCTGCCCAGCCAAGAACTTTTGGCCAAC 127918

Qy     301 ggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttc 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  127917 GGAAAATGCGATGATGGTCTCAAAGAGGAGCTCGCCCGGTTCAAACCCTATATATACTTC 127858

Qy     361 acatacatggacctacaggcacgagtaccactcaccgatatggctattatcagagatgag 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  127857 ACATACATGGACCTACAGGCACGAGTACCACTCACCGATATGGCTATTATCAGAGATGAG 127798
```

```
Qy     421 ttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggtt 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127797 TTCTTTGATTGTGATGAAATGATTTACATCAGCGACTTTAACTATTTCTTAGAAGAGGTT 127738

Qy     481 agcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaatttttgg 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127737 AGCATTCAAATCTTCAATAAATGTGGACTTATAGTTTGGTTCCAGGATGGAAAATTTTGG 127678

Qy     541 gttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaactttgat 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127677 GTTACCGTTGACCTCTTCATGGAAAAACAGGACTACTGCTTGTACCGGCACAACTTTGAT 127618

Qy     601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127617 TCGGATTTTCCAAAGTCCATGTGGATAATACGACACCGCTGCACAAGCCACATATCTCCT 127558

Qy     661 ggatccttagaga---------------------------------------------- 673
           ||||||||||||
Db 127557 GGATCCTTAGAGAGTAATTCAATTATTATTTATAACAAAAGTCATCAATATCGTTTTCAT 127498

Qy     674 ---------------------------------------------------------- 673

Db 127497 AGCTATATAGAAAGAAACGAAAATAACTCCTTATTGGTGTGCAAACTGCTGACTGCATTC 127438

Qy     674 ---------------------------------------------------------- 673

Db 127437 TGTGTGCCCTTTGCAAGTTTTTTAATTACAATAAGCTGGTTTTTTATTTCTTATATTTTC 127378

Qy     674 --------ttctaattataacaatgatatgctttgtcctaacaatcgcagtatatctata 725
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127377 TTTTTCAGTTCTAATTATAACAATGATATGCTTTGTCCTAACAATCGCAGTATATCTATA 127318

Qy     726 cattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcgaggtt 785
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127317 CATTAAGAAGCTGCGAAATGTTACTGGCAAGTGCATTGTATGCTGTATAGTTTCGAGGTT 127258

Qy     786 tatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgctctcc 845
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127257 TATCCAGTGCTTGATCATGATACTAGATCATTTAAATCTATTGAATGGCATTTGCTCTCC 127198

Qy     846 agct------------------------------------------------------- 849
           ||||
Db 127197 AGCTGGTTAGTTCGAATTGGAGATATCTCTGGAAGCTTTAAATGTTATATCGTTGACTTC 127138

Qy     850 -----ggttacagctcgcacttttttccggatggcttccaacctctggctctccgtcatca 904
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127137 TTGCAGGTTACAGCTCGCACTTTTTCCGGATGGCTTCCAACCTCTGGCTCTCCGTCATCA 127078

Qy     905 gctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaactatcggttcc 964
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127077 GCTACCATACGTGGAAAGTCTTGACGTCGCTCAATCGAGTCGACCCTAACTATCGGTTCC 127018

Qy     965 tgcggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttatatag 1024
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127017 TGCGGTACAACGCCTTCGTCTGGAGCACAGCCGCAATCATGACGGGAAGTATTTATATAG 126958

Qy    1025 ttaatcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggtttta 1084
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 126957 TTAATCAGATTTGGGAAAACGATCCCAGTAAATGGAACTGGTTGCCTCTGGTCGGTTTTA 126898

Qy    1085 ttcggtgctcggtcaa------------------------------------------- 1100
           ||||||||||||||||
Db 126897 TTCGGTGCTCGGTCAAAGGTAAACTTTCAATGCGATTGCTTTTAAATCGTTACCATGGTT 126838

Qy    1101 --------------------agattggcacccatccgtctggatctatataagtggaccg 1140
                               ||||||||||||||||||||||||||||||||||||||||
Db 126837 ACGTATTGCTTTTCCATATCAGATTGGCACCCATCCGTCTGGATCTATATAAGTGGACCG 126778

Qy    1141 tcgctggccctgagcactttcaatgtcgccatgttcgccctgacagccatttacattagg 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 126777 TCGCTGGCCCTGAGCACTTTCAATGTCGCCATGTTCGCCCTGACAGCCATTTACATTAGG 126718

Qy    1201 aaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggattaactgcata 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 126717 AAAGTGAAGGGGGGTATAAATAAGTTCACAAATGAGGAGGAGGGAAGGATTAACTGCATA 126658

Qy    1261 aactttgacagccagacgtaagatatgataaaaccttggaccttggtaaactagttatcc 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 126657 AACTTTGACAGCCAGACGTAAGATATGATAAAACCTTGGACCTTGGTAAACTAGTTATCC 126598

Qy    1321 ttgccgtattatccacatagttacctacagttcctgcggctctccatcgtgatgggcctt 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 126597 TTGCCGTATTATCCACATAGTTACCTACAGTTCCTGCGGCTCTCCATCGTGATGGGCCTT 126538

Qy    1381 acttggatattcaatgtcattccgtattctgcacggctccacattttctgggaatgggtc 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 126537 ACTTGGATATTCAATGTCATTCCGTATTCTGCACGGCTCCACATTTTCTGGGAATGGGTC 126478

Qy    1441 gggataatatccgagtattttcacagcgcgtttggaattgttttgttcgttctgctcgtc 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 126477 GGGATAATATCCGAGTATTTTCACAGCGCGTTTGGAATTGTTTTGTTCGTTCTGCTCGTC 126418

Qy    1501 ctgaagcgcagcacatggactctaatgatggattct 1536
           |||||||||||||||||||||||||||||||||||
Db 126417 CTGAAGCGCAGCACATGGACTCTAATGATGGATTCT 126382


RESULT   3
AC023733
LOCUS       AC023733    46913 bp    DNA             HTG       18-FEB-2000
DEFINITION  Drosophila melanogaster clone RP98-16J21, *** SEQUENCING IN
            PROGRESS ***, 41 unordered pieces.
ACCESSION   AC023733
VERSION     AC023733.2  GI:6996821
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 46913)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
            Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
            Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
            Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
            Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
            Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
            Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M.,
            Watlington,S., Weinstock,G., Weinstock,I.R., Williamson,A.,
            Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 46913)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-FEB-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Feb 18, 2000 this sequence version replaced gi:6984307.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: DRJW
```

```
    Center clone name: RP98-16J21
-------------- Summary Statistics
    Sequencing vector: M13; L08821
    Chemistry: Dye-terminator Big Dye: 99% of reads
    Assembly program: Phrap; version 0.980611
    Consensus quality: 15493 bases at least Q40
    Consensus quality: 25427 bases at least Q30
    Consensus quality: 31480 bases at least Q20
    Estimated insert size: 29542; sum-of-contigs estimation
    Quality coverage: 0.5x in Q20 bases; sum-of-contigs estimation
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 41 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1      755: contig of 755 bp in length
*      756      775: gap of unknown length
*      776     1598: contig of 823 bp in length
*     1599     1618: gap of unknown length
*     1619     3169: contig of 1551 bp in length
*     3170     3189: gap of unknown length
*     3190     3996: contig of 807 bp in length
*     3997     4016: gap of unknown length
*     4017     4859: contig of 843 bp in length
*     4860     4879: gap of unknown length
*     4880     5855: contig of 976 bp in length
*     5856     5875: gap of unknown length
*     5876     6968: contig of 1093 bp in length
*     6969     6988: gap of unknown length
*     6989     7947: contig of 959 bp in length
*     7948     7967: gap of unknown length
*     7968     8736: contig of 769 bp in length
*     8737     8756: gap of unknown length
*     8757     9500: contig of 744 bp in length
*     9501     9520: gap of unknown length
*     9521    10503: contig of 983 bp in length
*    10504    10523: gap of unknown length
*    10524    11456: contig of 933 bp in length
*    11457    11476: gap of unknown length
*    11477    12258: contig of 782 bp in length
*    12259    12278: gap of unknown length
*    12279    13525: contig of 1247 bp in length
*    13526    13545: gap of unknown length
*    13546    14301: contig of 756 bp in length
*    14302    14321: gap of unknown length
*    14322    15585: contig of 1264 bp in length
*    15586    15605: gap of unknown length
*    15606    16413: contig of 808 bp in length
*    16414    16433: gap of unknown length
*    16434    17352: contig of 919 bp in length
*    17353    17372: gap of unknown length
*    17373    18409: contig of 1037 bp in length
*    18410    18429: gap of unknown length
*    18430    19504: contig of 1075 bp in length
*    19505    19524: gap of unknown length
*    19525    20625: contig of 1101 bp in length
*    20626    20645: gap of unknown length
*    20646    21019: contig of 374 bp in length
*    21020    21039: gap of unknown length
*    21040    21850: contig of 811 bp in length
*    21851    21870: gap of unknown length
*    21871    22700: contig of 830 bp in length
*    22701    22720: gap of unknown length
*    22721    23720: contig of 1000 bp in length
*    23721    23740: gap of unknown length
*    23741    25116: contig of 1376 bp in length
*    25117    25136: gap of unknown length
*    25137    26853: contig of 1717 bp in length
*    26854    26873: gap of unknown length
*    26874    27573: contig of 700 bp in length
*    27574    27593: gap of unknown length
*    27594    28820: contig of 1227 bp in length
*    28821    28840: gap of unknown length
*    28841    29818: contig of 978 bp in length
*    29819    29838: gap of unknown length
*    29839    31525: contig of 1687 bp in length
*    31526    31545: gap of unknown length
*    31546    32670: contig of 1125 bp in length
*    32671    32690: gap of unknown length
*    32691    33781: contig of 1091 bp in length
*    33782    33801: gap of unknown length
*    33802    35134: contig of 1333 bp in length
*    35135    35154: gap of unknown length
*    35155    37141: contig of 1987 bp in length
*    37142    37161: gap of unknown length
*    37162    38615: contig of 1454 bp in length
*    38616    38635: gap of unknown length
*    38636    40858: contig of 2223 bp in length
*    40859    40878: gap of unknown length
*    40879    42886: contig of 2008 bp in length
*    42887    42906: gap of unknown length
*    42907    44311: contig of 1405 bp in length
*    44312    44331: gap of unknown length
*    44332    45841: contig of 1510 bp in length
*    45842    45861: gap of unknown length
*    45862    46913: contig of 1052 bp in length.
FEATURES             Location/Qualifiers
     source          1. .46913
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RP98-16J21"
BASE COUNT    12519 a  10639 c  10573 g  12323 t    859 others
ORIGIN


  Query Match           55.1%;  Score 846.6;  DB 68;  Length 46913;
  Best Local Similarity   79.2%;  Pred. No. 3.6e-225;
  Matches 1246;  Conservative    0;  Mismatches   19;  Indels  308;  Gaps    8;

Qy      1 atgaggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29953 ATGAGGCTTCCTTGGGTGATTTTCTGTACAGTTCTGCTTTTAATTTTTACGAACAACTCA 30012

Qy     61 aatgcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaagg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30013 AATGCAGATATTCCCGGGTGCAACTACTACGACACGGTTGATATCTCATACATTGAAAGG 30072

Qy    121 caaaacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactac 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30073 CAAAACGATTCGTATTTATATGATGACATCGAAATTCCTGCTAGCTTAACTGGATACTAC 30132

Qy    181 gagttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgt 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30133 GAGTTCAGGCAGTTTGGGGACGGTTCGATTACGCCGATTGAAAAGCATTTAAGGGCCTGT 30192

Qy    241 gtttgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttttggccaac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30193 GTTTGCAGCGTGCGGCCCTGTATTCGAATCTGCTGCCCAGCCAAGAACTTTTTGGCCAAC 30252

Qy    301 ggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30253 GGAAAATGCGATGATGGTCTCAAAGAGGAGCTCGCCCGGTTCAAACCCTATATATACTTC 30312

Qy    361 acatacatggacctacaggcacgagtaccactcaccgatatggctattatcagagatgag 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30313 ACATACATGGACCTACAGGCACGAGTACCACTCACCGATATGGCTATTATCAGAGATGAG 30372

Qy    421 ttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggtt 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30373 TTCTTTGATTGTGATGAAATGATTTACATCAGCGACTTTAACTATTTCTTAGAAGAGGTT 30432

Qy    481 agcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaatttgg 540
```

```
Db  30433 AGCATTCAAATCTTCAATAAATGTGGACTTATAGTTTGGTTCCAGGATGGAAAATTTTGG 30492

Qy    541 gttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaactttgat 600
Db  30493 GTTACCGTTGACCTCTTCATGGAAAAACAGGACTACTGCTTGTACCGGCACAACTTTGAT 30552

Qy    601 tcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacatatctcct 660
Db  30553 TCGGATTTTCCAAAGTCCATGTGGATAATACGACACCGCTGCACAAGCCACATATCTCCT 30612

Qy    661 ggatccttagaga---------------------------------------------- 673
Db  30613 GGATCCTTAGAGAGTAATTCAATTATTATTTATAACAAAAGTCATCAATATCGTTTTCAT 30672

Qy    674 ---------------------------------------------------------- 673
Db  30673 AGCTATATAGAAAGAAACGAAAATAACTCCTTATTGGTGTGCAAACTGCTGACTGCATTC 30732

Qy    674 ---------------------------------------------------------- 673
Db  30733 TGTGTGCCCTTTGCAAGTTTTTTAATAACAATAAGCTGGTTTTTTATTTCTTATATTTTC 30792

Qy    674 --------ttctaattataacaatgatatgctttgtcctaacaatcgcagtatatctata 725
Db  30793 TTTTTCAGTTCTAATTATAACAATGATATGCTTTGTCCTAACAATCGCAGTATATCTATA 30852

Qy    726 cattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcgaggtt 785
Db  30853 CATTAAGAAGCTGCGAAATGTTACTGGCAAGTGCATTGTATGCTGTATAGTTTCGAGGTT 30912

Qy    786 tatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgctctcc 845
Db  30913 TATCCAGTGCTTGATCATGATACTAGATCATTTAAATCTATTGAATGGCATTTGCTCTCC 30972

Qy    846 agct------------------------------------------------------ 849
Db  30973 AGCTGGTTAGTTTTTATTGGAGATATCTCTGGAAGCTTTAAATGTTATATCGTTGACTTC 31032

Qy    850 -----ggttacagctcgcacttttttccggatggcttccaacctctggctctccgtcatca 904
Db  31033 TTGCAGGTTACAGCTCGCACTTTTTCCGGATGGCTTCCAACCTCTGGCTCTCCGTCATCA 31092

Qy    905 gctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaactatcggttcc 964
Db  31093 GCTACCATACGTGGAAAGTCTTGACGTCGCTCAATCGAGTCGACCCTAACTATCGGTTCC 31152

Qy    965 tgcggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttatatag 1024
Db  31153 TGCGGTACAACGCCTTCGTCTGGAGCACAGCCGCAATCATGACGGGAAGTATTTATATAG 31212

Qy   1025 ttaatcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggtttta 1084
Db  31213 TTAATCAGATTTGGGAAAACGATCCCAGTAAATGGAACTGGTTGCCTCTGGTCGGTTTTA 31272

Qy   1085 ttcggtgctcggtcaa---------------------------------------------- 1100
Db  31273 TTCGGTGCTCGGTCAAAGGTAAACTNTCAATGCGATTGCTTTTAAATCGTTACCATGGTT 31332

Qy   1101 -------------------agattggcacccat-ccgtctggatctatataagtggacc 1139
Db  31333 ACGTATTGCTTTTCCATATCAGATTGGCACCCATCCCGTCTGGATCTATATAAGTGGGAC 31392

Qy   1140 gtcgc---tggccctgagcacttt--caatgtcgccatgttc-gccctgacagccattta 1193
Db  31393 CGTCCGCTGGCCCCTGAGCACTTTTCAATGTCCGCCATGTTCGGCCCTGACAGCCATTTA 31452

Qy   1194 catta-ggaaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggatta 1252
Db  31453 CATTAGGGAAGGTGAGGGGGGGTATTAATAAGTTCACAAATGAGGAAGAGGGAAGGATTA 31512

Qy   1253 actgcataaactt 1265
Db  31513 ACTGCCTAACCTT 31525
```

```
RESULT   4
AF053725
LOCUS       AF053725     4616 bp    DNA             INV       31-MAR-1998
DEFINITION  Drosophila melanogaster myristoyl-CoA: protein N-myristoyl
            transferase (Nmt) gene, complete cds.
ACCESSION   AF053725
VERSION     AF053725.1  GI:2996201
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 4616)
  AUTHORS   Ntwasa,M., Egerton,M. and Gay,N.J.
  TITLE     Sequence and expression of Drosophila myristoyl-CoA: protein
            N-myristoyl transferase: evidence for proteolytic processing and
            membrane localisation
  JOURNAL   J. Cell. Sci. 110 (Pt 2), 149-156 (1997)
  MEDLINE   97196956
REFERENCE   2  (bases 1 to 4616)
  AUTHORS   Ntwasa,M., Egerton,M. and Gay,N.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-1998) Anatomy and Cell Biology, University of
            Cape Town, Anzio Road, Observatory, Cape Town 7925, South Africa
FEATURES             Location/Qualifiers
     source          1. .4616
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="3"
                     /map="3L; 66C"
     gene            <1729. .>2892
                     /gene="Nmt"
     CDS             1729. .2892
                     /gene="Nmt"
                     /note="dNMT"
                     /codon_start=1
                     /product="myristoyl-CoA: protein N-myristoyl transferase"
                     /protein_id="AAC08578.1"
                     /db_xref="GI:2996202"
                     /translation="MASTRQMAKKFAFWSTQPVTKLDEQVTTNECIEPNKEISEIRAL
                     PYTLPGFKWVTLDLNDANDLKELYTLLNENYVEDDDAMFRFDYQPEFLKWSLQPPGWK
                     RDWHVGVRVEKSGKLVGFISAIPSKLKSYDKVLKVVDINFLCVHKKLRSKRVAPVLIR
                     EISRRVNLTGIFQAAYTAGVVLPTPVATCRYWHRSLNPKKLVDVRFSHLARNMTMQRT
                     MKLYKLPDQPKTKGYRRITAKDMDKAHKLLEDYLKRFQLSPVFSKEEFRHWFTPKEGI
                     IDCFVVADEKGNITDLTSYYCLPSSVMHHPVHKTVRAAYSFYNVSTKTPWLDLMGDAL
                     ISARNVQMDVYNALDLMENKKYFAPLKFGAGDGNLQYYLYNWRCPSMQPERLL"
BASE COUNT     1297 a   1064 c   1048 g   1207 t
ORIGIN
```

```
  Query Match             29.9%;  Score 460;  DB 5;  Length 4616;
  Best Local Similarity   87.3%;  Pred. No. 2.4e-117;
  Matches  578;  Conservative    0;  Mismatches   15;  Indels   69;  Gaps    4;

Qy    944 tcgaccctaactatcggttcctgcggtacaacgccttcgtctggagcacagccgcaatca 1003
Db      1 TCGACCCTAACTATCGGTTCCTGCGGTACAACGCCTTCGTCTGGAGCACAGCCGCAATCA 60

Qy   1004 tgacgggaagtatttatatagttaatcagatttgggaaaacgatcccagtaaatggaact 1063
Db     61 TGACGGGAAGTATTTATATAGTTAATCAGATTTGGGAAAACGATCCCAGTAAATGGAACT 120

Qy   1064 ggttgcctctggtcggttttattcggtgctcggtcaa---------------------- 1100
Db    121 GGTTGCCTCTGGTCGGTTTTATTCGGTGCTCGGTCAAAGGTAAACTTTCAATGCGATTGC 180

Qy   1101 ----------------------------------------agattggcacccatccgtc 1119
Db    181 TTTTAAATCGTTACCATGGTTACGTATTGCTTTTCCATATCAGATTGGCACCCATCCGTC 240
```

```
Qy   1120 tggatctatataagtggaccgtcgctggccctgagcactttcaatgtcgccatgttcgcc 1179
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGGATCTATATAAGCGGACCGTCGCTGGCCCTGAGCACTTTCAATGTCGCCATGTTCGCC 300

Qy   1180 ctgacagccatttacattaggaaagtgaaggggggtataaataagttcacaaatgaggag 1239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTGACAGCCATTTACATTAGGAAAGTGAAGGGGGGTATAAATAAGTTCACAAATGAGGAG 360

Qy   1240 gagggaaggattaactgcataaactttgacagccagacgtaagatatgataaaaccttgg 1299
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    361 GAGGGAAGGATTAACTGCATCAACTTTGACAGCCAGACGTAAGATATGATAAAACCTTGG 420

Qy   1300 accttggtaaactagttatccttgccgtattatccacatagttacctacagttcctgcgg 1359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ACCTTGGTAAACTAGTTATCCTTGCCGTATTATCCACATAGTTACCTACAGTTCCTGCGG 480

Qy   1360 ctctccatcgtgatggg-ccttacttggatattcaatgtcattcc---gtattctgcacg 1415
          |||||||||||||||| |||||||||||||||||||||||| |   |  | |  ||
Db    481 CTCTCCATCGTGATGGGTCCTTACTTGGATATTCAATGTCATTTCCGTATTCTGTCATCG 540

Qy   1416 gctccacattttctgggaatggg-tcgggataatatccgagtattttcacagcgcgtttg 1474
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    541 GCTCCACATTTTCTGGGAATGGGTTCGGGATAATATCCGAGTATTTTCACAGCGCGTTTG 600

Qy   1475 gaattgttttgttcgttctgctcgtcctgaagcgcagcacatggactctaatgatggatt 1534
          |||| |||||| ||| |||||||| ||||||| ||||||||||||||||||||||||||
Db    601 GAATCGTTTTGATCGATCTGCTCGACCTGAAGGGCAGCACATGGACTCTAATGATGGATT 660

Qy   1535 ct 1536
          ||
Db    661 CT 662


RESULT   5
AC010037/c
LOCUS       AC010037   164659 bp    DNA             HTG       29-MAY-2000
DEFINITION  Drosophila melanogaster clone RPCI98-6E4, WORKING DRAFT SEQUENCE,
            46 unordered pieces.
ACCESSION   AC010037
VERSION     AC010037.5  GI:8101304
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 164659)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
            Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
            Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
            Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
            Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
            Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
            Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M.,
            Watlington,S., Weinstock,G., Weinstock,I.R., Williamson,A.,
            Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 164659)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1999) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 29, 2000 this sequence version replaced gi:6996724.
            ---------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            ---------- Project Information
                Center project name: DRCO
                Center clone name: RPCI98-6E4
            ---------- Summary Statistics
                Sequencing vector: M13; L08821
                Chemistry: Dye-primer Bodipy: 29% of reads
                Chemistry: Dye-terminator Big Dye: 25% of reads
                Assembly program: Phrap; version 0.980611
                Consensus quality: 112955 bases at least Q40
                Consensus quality: 133293 bases at least Q30
                Consensus quality: 144574 bases at least Q20
                Estimated insert size: 160113; sum-of-contigs estimation
                Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
            ----------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 46 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     6150: contig of 6150 bp in length
            *     6151     6250: gap of unknown length
            *     6251    13075: contig of 6825 bp in length
            *    13076    13175: gap of unknown length
            *    13176    18980: contig of 5805 bp in length
            *    18981    19080: gap of unknown length
            *    19081    24693: contig of 5613 bp in length
            *    24694    24793: gap of unknown length
            *    24794    33330: contig of 8537 bp in length
            *    33331    33430: gap of unknown length
            *    33431    40699: contig of 7269 bp in length
            *    40700    40799: gap of unknown length
            *    40800    46106: contig of 5307 bp in length
            *    46107    46206: gap of unknown length
            *    46207    50498: contig of 4292 bp in length
            *    50499    50598: gap of unknown length
            *    50599    55756: contig of 5158 bp in length
            *    55757    55856: gap of unknown length
            *    55857    59777: contig of 3921 bp in length
            *    59778    59877: gap of unknown length
            *    59878    66349: contig of 6472 bp in length
            *    66350    66449: gap of unknown length
            *    66450    71183: contig of 4734 bp in length
            *    71184    71283: gap of unknown length
            *    71284    76459: contig of 5176 bp in length
            *    76460    76559: gap of unknown length
            *    76560    81902: contig of 5343 bp in length
            *    81903    82002: gap of unknown length
            *    82003    86768: contig of 4766 bp in length
            *    86769    86868: gap of unknown length
            *    86869    92357: contig of 5489 bp in length
            *    92358    92457: gap of unknown length
            *    92458    97795: contig of 5338 bp in length
            *    97796    97895: gap of unknown length
            *    97896   101452: contig of 3557 bp in length
            *   101453   101552: gap of unknown length
            *   101553   105177: contig of 3625 bp in length
            *   105178   105277: gap of unknown length
            *   105278   108706: contig of 3429 bp in length
            *   108707   108806: gap of unknown length
            *   108807   111390: contig of 2584 bp in length
            *   111391   111490: gap of unknown length
            *   111491   113983: contig of 2493 bp in length
```

```
          *   113984   114083: gap of unknown length
          *   114084   116768: contig of 2685 bp in length
          *   116769   116868: gap of unknown length
          *   116869   119023: contig of 2155 bp in length
          *   119024   119123: gap of unknown length
          *   119124   121906: contig of 2783 bp in length
          *   121907   122006: gap of unknown length
          *   122007   124731: contig of 2725 bp in length
          *   124732   124831: gap of unknown length
          *   124832   126964: contig of 2133 bp in length
          *   126965   127064: gap of unknown length
          *   127065   129583: contig of 2519 bp in length
          *   129584   129683: gap of unknown length
          *   129684   132383: contig of 2700 bp in length
          *   132384   132483: gap of unknown length
          *   132484   134455: contig of 1972 bp in length
          *   134456   134555: gap of unknown length
          *   134556   136568: contig of 2013 bp in length
          *   136569   136668: gap of unknown length
          *   136669   138348: contig of 1680 bp in length
          *   138349   138448: gap of unknown length
          *   138449   141263: contig of 2815 bp in length
          *   141264   141363: gap of unknown length
          *   141364   143990: contig of 2627 bp in length
          *   143991   144090: gap of unknown length
          *   144091   146194: contig of 2104 bp in length
          *   146195   146294: gap of unknown length
          *   146295   148427: contig of 2133 bp in length
          *   148428   148527: gap of unknown length
          *   148528   150406: contig of 1879 bp in length
          *   150407   150506: gap of unknown length
          *   150507   152893: contig of 2387 bp in length
          *   152894   152993: gap of unknown length
          *   152994   154351: contig of 1358 bp in length
          *   154352   154451: gap of unknown length
          *   154452   156305: contig of 1854 bp in length
          *   156306   156405: gap of unknown length
          *   156406   157843: contig of 1438 bp in length
          *   157844   157943: gap of unknown length
          *   157944   159199: contig of 1256 bp in length
          *   159200   159299: gap of unknown length
          *   159300   160524: contig of 1225 bp in length
          *   160525   160624: gap of unknown length
          *   160625   161884: contig of 1260 bp in length
          *   161885   161984: gap of unknown length
          *   161985   163299: contig of 1315 bp in length
          *   163300   163399: gap of unknown length
          *   163400   164659: contig of 1260 bp in length.
FEATURES             Location/Qualifiers
     source          1. .164659
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RPCI98-6E4"
     misc_feature    complement(16179. .16278)
                     /note="assembly_fragment:contig14;
                     vector:Vector_pBACe3.6;
                     vector_side:right"
     misc_feature    complement(53461. .53560)
                     /note="assembly_fragment:contig29;
                     vector:Vector_pBACe3.6;
                     vector_side:left"
BASE COUNT    44679 a  34729 c  35141 g  45537 t   4573 others
ORIGIN


  Query Match           28.4%;  Score 436;  DB 61;  Length 164659;
  Best Local Similarity  100.0%;  Pred. No. 2e-110;
  Matches  436;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1101 agattggcacccatccgtctggatctatataagtggaccgtcgctggccctgagcacttt 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  59662 AGATTGGCACCCATCCGTCTGGATCTATATAAGTGGACCGTCGCTGGCCCTGAGCACTTT 59603

Qy   1161 caatgtcgccatgttcgccctgacagccatttacattaggaaagtgaaggggggtataaa 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  59602 CAATGTCGCCATGTTCGCCCTGACAGCCATTTACATTAGGAAAGTGAAGGGGGGTATAAA 59543

Qy   1221 taagttcacaaatgaggaggagggaaggattaactgcataaactttgacagccagacgta 1280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  59542 TAAGTTCACAAATGAGGAGGAGGGAAGGATTAACTGCATAAACTTTGACAGCCAGACGTA 59483

Qy   1281 agatatgataaaaccttggaccttggtaaactagttatccttgccgtattatccacatag 1340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  59482 AGATATGATAAAACCTTGGACCTTGGTAAACTAGTTATCCTTGCCGTATTATCCACATAG 59423

Qy   1341 ttacctacagttcctgcggctctccatcgtgatgggccttacttggatattcaatgtcat 1400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  59422 TTACCTACAGTTCCTGCGGCTCTCCATCGTGATGGGCCTTACTTGGATATTCAATGTCAT 59363

Qy   1401 tccgtattctgcacggctccacattttctgggaatgggtcgggataatatccgagtattt 1460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  59362 TCCGTATTCTGCACGGCTCCACATTTTCTGGGAATGGGTCGGGATAATATCCGAGTATTT 59303

Qy   1461 tcacagcgcgtttggaattgttttgttcgttctgctcgtcctgaagcgcagcacatggac 1520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  59302 TCACAGCGCGTTTGGAATTGTTTTGTTCGTTCTGCTCGTCCTGAAGCGCAGCACATGGAC 59243

Qy   1521 tctaatgatggattct 1536
          ||||||||||||||||
Db  59242 TCTAATGATGGATTCT 59227


RESULT   6
AC017388
LOCUS       AC017388    42611 bp    DNA             HTG       09-DEC-1999
DEFINITION  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
            pieces.
ACCESSION   AC017388
VERSION     AC017388.1  GI:6553598
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 42611)
  AUTHORS   Adams,M. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     This sequence was identified as CDM:10210530 by the submitter.
            For more information on this record e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .42611
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
BASE COUNT    12411 a   8643 c   8940 g  12617 t
ORIGIN


  Query Match           12.8%;  Score 196;  DB 64;  Length 42611;
  Best Local Similarity  52.0%;  Pred. No. 1.8e-43;
  Matches  634;  Conservative    0;  Mismatches  470;  Indels  116;  Gaps    4;

Qy      4 aggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactcaaat 63
          ||  | | |  |   |||||  || | ||||  | ||  |  || | | | | ||  |
Db  29741 AGAATGCTTCTGAATATTTTAGCTATAATTCTTGTATTCGTAATTTCAAGCCAATCGGAG 29800

Qy     64 gcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaaggcaa 123
          |||| |||||| || ||  | |||| |||||| || ||||| ||  | ||   || || |
Db  29801 GCAGTTATTCCTGGATGTGATTACTTCGACACCGTGGATATATCTCATATCCCAAAGCTA 29860
```

```
Qy    124 aacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactacgag 183
          || || || |||   || || ||  |    ||||||||     | || ||    ||
Db  29861 AATGACTCCTATGCGTACGAGGAGCTAATCATTCCTGCTCATCTCACGGGTCTATATACA 29920

Qy    184 ttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgtgtt 243
          || |||||  | | ||| || ||      |||  |  | |  || ||||| || ||| |
Db  29921 TTTAGGCAACTGGCGGATGGATCACAGGAGCCAGTGAAGAGTCACTTAAGAGCATGTATC 29980

Qy    244 tgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttggccaacgga 303
          ||||  |    || ||||| ||| ||||||||||   ||| ||  |  || | ||  |
Db  29981 TGCAAACTAAAACCTTGTATCCGATTCTGCTGCCCCCGCAACAAAATGATGCCGAATAGT 30040

Qy    304 aaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttcaca 363
             ||    ||||| |||| |||| | ||    ||| |||| |||||  |  |  |||||
Db  30041 CGCTGTAGCGATGGACTCACAGAGAATCTGAAGCGGATCAATCCCTACTTGAAGATCACA 30100

Qy    364 tacatggacctacaggcacgagta------ccactcaccgatatggctattatcagagat 417
              |||   |     | ||  |       |  |||||||| |||    |  | ||| ||
Db  30101 CTGGAGGATGGAACTATAGGAAAATATTACCTTCTCACCGACATGATCGTCTTAAGATAT 30160

Qy    418 gagttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagag 477
          |||||     |||| || |  |  |||    |    ||   | | ||     || | |||
Db  30161 GAGTTTCGGTATTGCGAGAAGGTTGTTTCGGTACAGGAGGATCAGTACAAGCTATATGAG 30220

Qy    478 gttagcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaattt 537
          ||||  |||  ||||||    || | |     |||| |||| |  |   |
Db  30221 GTTAAGATTTCAATCTTTTTGAAGTATCTCAGTATAATTTGATATCTATAGAATGGCAGC 30280

Qy    538 tgggttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaacttt 597
          |   | |       |    | |     | |  |  | |    | |||| | | | |
Db  30281 TTTATGATTAAACCCGATGTAAACTGGACATTGAGCAAGCAGTGGTACTGCCTCCATCCC 30340

Qy    598 gattcggattttccaaagtccatgtggataatacgaca---------------------- 635
            || |||   ||| || ||||| ||||||  |   |||
Db  30341 CGTTTGGAAGATCCGAATTCCATATGGATACTGGAACATGTATATATACCGAAGTCAATG 30400

Qy    636 ----ccgctgcacaagccacatatctcctggatccttagagattct-------------- 677
              |||   | || | ||   | || ||  ||| |   |  |  |
Db  30401 CCAGCCGTACCTCAGGGCAAGCACCTACTATATCATCTAATCTCATCTCATCTAATTCTC 30460

Qy    678 ----------------aattataacaatgatatgctttgtcctaacaatcgcagtatatc 721
                         | |||  | ||| |  | | | | || || || ||||| ||||
Db  30461 TGATTTTTCCAGTCGGTACTATTTCCATGGTTGGTTGTATTCTCACGATTGCAGTGTATC 30520

Qy    722 tatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcga 781
          | ||||| ||||| |||||||||| |   ||| || ||  |  | |||| | || |||  |
Db  30521 TCTACATAAAGAAACTGCGAAATTTGCTTGGAAAATGTTTCATTTGCTATGTATTTTGTA 30580

Qy    782 ggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgct 841
           |||| | |||| | | |||  |  |  ||   |||| ||  | | ||||   |||||||
Db  30581 AGTTTGTGCAGTACCTTATCTGGGCAGGAGGAGATTTGAACTTGTGGAATAATATTTGCT 30640

Qy    842 ctccag---------------------------------------------------- 847
          | | ||
Db  30641 CCCTAGCAGGTTAGTTGAATCCATATTAAAGTTTGACTTAAAATATGCGTTTATTCTAAA 30700

Qy    848 ctggttacagctcgcacttttttccggatggcttccaacctctggctctccgtcatcagct 907
          | |||||||      | || ||     |||||||  || | |||||||| || || ||
Db  30701 CAGGTTACACTAACTATTTCTTTGCTTTGGCTTCTCACTTTTGGCTCTCTGTTATGAGTC 30760

Qy    908 accatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaactatcggttcctgc 967
          | || |  ||||||  ||||| ||   |||||||  |  || | || |||||  || ||
Db  30761 ATCAAATATGGAAAAACTTGAGGTTAATCAATCGGGATGAACGTAGCTATCACTTTCTCA 30820

Qy    968 ggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttatatagtta 1027
            ||||||  || ||  ||| ||||| | || ||||||||||| || ||  ||| | ||
Db  30821 TCTACAACATCTACGGTTGGGGCACACCTGCTATCATGACGGCAATTACATATCTGGTGG 30880

Qy   1028 atcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggttttattc 1087
          ||  |   |||||  |    ||    |||  ||| ||| | ||    || || ||   ||
Db  30881 ATTGGGCCTGGGAGGATAGACCTGACAAACTGAATTGGATTCCCGGTGTTGGCTTATATC 30940

Qy   1088 ggtgctcggtcaaagattgg 1107
```

```
          | ||||  | | ||   | ||
Db  30941 GATGCTGGATAAACAGTGGG 30960

RESULT   7
AE003558
LOCUS       AE003558   273744 bp    DNA             INV       06-OCT-2000
DEFINITION  Drosophila melanogaster genomic scaffold 142000013386050 section 10
            of 54, complete sequence.
ACCESSION   AE003558 AE002602
VERSION     AE003558.2  GI:10728071
KEYWORDS    HTG.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 273744)
  AUTHORS   Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
            Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
            George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
            Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
            Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
            Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
            Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
            Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
            Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
            Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
            Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
            Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
            Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
            Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
            Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
            Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S.,
            Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S.,
            Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
            Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
            Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
            Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
            Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
            Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
            Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
            McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
            Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
            Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
            Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
            Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
            Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
            Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
            Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,
            Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
            Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,
            Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
            Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
            Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
            Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
  TITLE     The genome sequence of Drosophila melanogaster
  JOURNAL   Science 287 (5461), 2185-2195 (2000)
  MEDLINE   20196006
REFERENCE   2  (bases 1 to 273744)
  AUTHORS   Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     On Oct 9, 2000 this sequence version replaced gi:7295193.
FEATURES             Location/Qualifiers
     source          1. .273744
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="3L"
     mRNA            join(<6847. .7482,7587. .8054,8129. .8527)
                     /gene="CG16992"
                     /product="CT37715"
```

```
                     /db_xref="FLYBASE:FBan0016992"
                     /db_xref="FLYBASE:FBgn0035789"
     gene            <6847. .>8527
                     /gene="CG16992"
                     /map="66A1-66A1"
                     /db_xref="FLYBASE:FBan0016992"
                     /db_xref="FLYBASE:FBgn0035789"
     CDS             6847. .7365
                     /gene="CG16992"
                     /note="CG16992 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0016992"
                     /db_xref="FLYBASE:FBgn0035789"
                     /protein_id="AAF50551.1"
                     /db_xref="GI:7295228"
                     /translation="MLLNILAIILVFVISSQSEAVIPGCDYFDTVDISHIPKLNDSYA
                     YEELIIPAHLTGLYTFRQLADGSQEPVKSHLRACICKLKPCIRFCCPRNKMMPNSRCS
                     DGLTENLKRINPYLKITLEDGTIGKYYLLTDMIVLRYEFRYCEKVVSVQEDQYKLYEV
                     KISIFLKYLSII"
     mRNA            join(<12641. .12797,12847. .13276,13337. .13875,
                     13951. .>14261)
                     /gene="Cyp316a1"
                     /product="CT15730"
                     /db_xref="FLYBASE:FBan0008540"
                     /db_xref="FLYBASE:FBgn0035790"
     gene            <12641. .>14261
                     /gene="Cyp316a1"
                     /note="CG8540"
                     /map="66A1-66A2"
                     /db_xref="FLYBASE:FBan0008540"
                     /db_xref="FLYBASE:FBgn0035790"
     CDS             join(12641. .12797,12847. .13276,13337. .13875,13951. .14261)
                     /gene="Cyp316a1"
                     /note="Cyp316a1 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0008540"
                     /db_xref="FLYBASE:FBgn0035790"
                     /protein_id="AAF50550.2"
                     /db_xref="GI:10728072"
                     /translation="MILTATFICFCLASAFNYFRARRQRSLIKNLKGPFTWPLMGAMH
                     KLLFLTPINFFQRSTEYLTKYGTFSRCWVFHRLFIPLADLELSRQLLENDTHLETGYE
                     LMKDWLVGGVLMCQSEQWQKRHSLISGLFDKGNLEQLIDLSRHQTEQLLQKLAKQADQ
                     KVFDIWYTVSPIVLDLMVMTTCGAKPSEEYSKNLKDLSEIYRKRFLSLQSANRFNYWL
                     SSPFMRKRQNRLIKRLNDEHNNLMAMHQSQNQLKIENGLDIYQLRPIPLKDHKSLLEI
                     LLESKDPQLTGEEICGELNTCNYLGYQLCSPALCFCLVTIARNPSVQQKCLDELNLAQ
                     IKDQGWDLEKLNYLDAVLHETMRLYPPQVIVGRQLKKDFPYNAELPCGSEIYINLYEL
                     QRNEVRYPKANHFDAQRFLDSPPELLSYSLGPRCCPARKFSMQLLKTLLAPILANFEV
                     LPYGDEVRLDLRLVLGSSNGFQLALKPR"
     mRNA            complement(join(<14395. .14744,14805. .15361,15433. .15886,
                     16806. .>16962))
                     /gene="Cyp4d8"
                     /note="Nucleotide sequence of the Celera sequence differs
                     from the published sequence for this transcript."
                     /product="CT14135"
                     /db_xref="FLYBASE:FBan0004321"
                     /db_xref="FLYBASE:FBgn0015033"
     gene            complement(<14395. .>16962)
                     /gene="Cyp4d8"
                     /note="CG4321"
                     /map="66A1-66A2"
                     /db_xref="FLYBASE:FBan0004321"
                     /db_xref="FLYBASE:FBgn0015033"
     CDS             complement(join(14395. .14744,14805. .15361,15433. .15886,
                     16806. .16962))
                     /gene="Cyp4d8"
                     /note="Cyp4d8 gene product; Nucleotide sequence of the
                     Celera sequence differs from the published sequence for
                     this transcript"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0004321"
                     /db_xref="FLYBASE:FBgn0015033"
                     /protein_id="AAF50549.1"
                     /db_xref="GI:7295226"
                     /translation="MLLFLLVVLLFGAGWIIHLGQADRRRKVANLPGPICPPLIGAMQ
                     LMLRLNPKTFIKVGREYVLKFGHLQRVWIFNRLLIMSGDAELNEQLLSSQEHLVKHPV
                     YKVLGQWLGNGLLLSDGKVWHQRRKIITPTFHFSILEQFVEVFDQQSNICVQRLAQKA
                     NGNTFDVYRSICAAALDIIAETAMGTKIYAQANESTPYAEAVNECTALLSWRFMSVYL
                     QVELLFTLTHPHLKWRQTQLIRTMQEFTIKVIEKRRQALEDQQSKLMDTADEDVGSKR
                     RMALLDVLLMSTVDGRPLTNDEIREEVDTFMFEGHDTTTSALSFCLHELSRHPEVQAK
                     MLEEIVQVLGTDRSRPVSIRDLGELKYMECVIKESLRMYPPVPIVGRKLQTDFKYTHS
                     VHGDGVIPAGSEIIIGIFGVHRQPETFPNPDEFIPERHENGSRVAPFKMIPFSAGPRN
                     CIGQKFAQLEMKMMLAKIVREYELLPMGQRVECIVNIVLRSETGFQLGMRKRKHN"
     mRNA            complement(join(18046. .18614,18698. .19448))
                     /gene="CG8539"
                     /product="CT14125"
                     /db_xref="FLYBASE:FBan0008539"
                     /db_xref="FLYBASE:FBgn0035791"
     gene            complement(<18046. .>19448)
                     /gene="CG8539"
                     /map="66A1-66A2"
                     /db_xref="FLYBASE:FBan0008539"
                     /db_xref="FLYBASE:FBgn0035791"
     CDS             complement(join(18170. .18614,18698. .19410))
                     /gene="CG8539"
                     /note="CG8539 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0008539"
                     /db_xref="FLYBASE:FBgn0035791"
                     /protein_id="AAF50548.1"
                     /db_xref="GI:7295225"
                     /translation="MWKILCGIVVLVTIVVNQLSEAREVRRRRGLMLQLDNYLSYDGI
                     MQYLDELALSHSNRVTLKDVARTYENRALKMAIITNGDGRPGKRVIFLDAALHSREWM
                     TPAAALLTIHKLVVEFAENSDLLTDYDWHIMPLANPDGYEYSRNTERYWRNTRTPNGG
                     NCFGTNLNRNFAVDWNVGFPELKDPCDENYAGSSPFSEVEARTVRDIMHGLVESKRAV
                     MYLSLHTANRSVFYPWVYDTDPVSNQKEHDEIGRFVADRILQSTGTFIKTWQYAKYAG
                     TFGGTSMDYALLAGFPLSFVFEMSGTGRDHVEYKFFPPARDIRHLAEESWTGIKAFAE
                     KTIEKYPPSRVISYNPMVKLAENAATRNLGFNSMVATVCLFFTYLTNCNFF"
     mRNA            complement(<25839. .>26432)
                     /gene="CG7548"
                     /product="CT14117"
                     /db_xref="FLYBASE:FBan0007548"
                     /db_xref="FLYBASE:FBgn0035792"
     gene            complement(<25839. .>26432)
                     /gene="CG7548"
                     /map="66A1-66A2"
                     /db_xref="FLYBASE:FBan0007548"
                     /db_xref="FLYBASE:FBgn0035792"
     CDS             complement(25839. .26432)
                     /gene="CG7548"
                     /note="CG7548 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0007548"
                     /db_xref="FLYBASE:FBgn0035792"
                     /protein_id="AAF50547.1"
                     /db_xref="GI:7295224"
                     /translation="MAYFQFLSVASSLLLLLAPTWAGLIAQPTITYAGNEHAVAHTQQ
                     NVVRSFDGTVSHYAKSLATPYSQVHKQDTRISNNVYQPAVAKTFSYATAPAPVYTHQA
                     HQEPKNLFTQASPVYQQDAHVQTPSVYQQAAHVQTPSVYHQPAQVQSYHQPGHVEAPS
                     VYQQNAHYSHQPAVIHYSPAETVSHMSFDGFGTHYGF"
     mRNA            join(31242. .31673,31742. .31859,31936. .32147,32204. .32959,
                     33017. .33623,33754. .33852,34049. .34585,35279. .35812,
                     35872. .36008,36099. .36281,36347. .36777)
                     /gene="CG7546"

  Query Match          12.8%;  Score 196;  DB 4;  Length 273744;
  Best Local Similarity  52.0%;  Pred. No. 2.3e-43;
  Matches  634;  Conservative    0;  Mismatches  470;  Indels  116;  Gaps    4;

Qy      4 aggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactcaaat 63
          ||  | | |  |   |||||   || | ||||  | ||  |  || | | | | ||  |
Db   6844 AGAATGCTTCTGAATATTTTAGCTATAATTCTTGTATTCGTAATTTCAAGCCAATCGGAG 6903

Qy     64 gcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaaggcaa 123
          |||| |||||| || ||  | |||| |||||| || ||||| ||  | ||   || || |
Db   6904 GCAGTTATTCCTGGATGTGATTACTTCGACACCGTGGATATATCTCATATCCCAAAGCTA 6963
```

```
Qy    124 aacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactacgag 183
          || || || |||   || || ||  |    |||||||||    | || ||    ||
Db   6964 AATGACTCCTATGCGTACGAGGAGCTAATCATTCCTGCTCATCTCACGGGTCTATATACA 7023

Qy    184 ttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgtgtt 243
          || |||||  | | ||| || ||      |||  |  | |  || ||||| || ||| |
Db   7024 TTTAGGCAACTGGCGGATGGATCACAGGAGCCAGTGAAGAGTCACTTAAGAGCATGTATC 7083

Qy    244 tgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttggccaacgga 303
          ||||  |    || ||||| ||| ||||||||||   ||| ||  |  || | ||  |
Db   7084 TGCAAACTAAAACCTTGTATCCGATTCTGCTGCCCCCGCAACAAAATGATGCCGAATAGT 7143

Qy    304 aaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttcaca 363
             ||   ||||| |||| |||| | ||    ||| |||| |||||  |  |  |||||
Db   7144 CGCTGTAGCGATGGACTCACAGAGAATCTGAAGCGGATCAATCCCTACTTGAAGATCACA 7203

Qy    364 tacatggacctacaggcacgagta------ccactcaccgatatggctattatcagagat 417
             |||  |     | ||  |      |  |||||||| |||    |  | ||| ||
Db   7204 CTGGAGGATGGAACTATAGGAAAATATTACCTTCTCACCGACATGATCGTCTTAAGATAT 7263

Qy    418 gagttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagag 477
          |||||     |||| ||  |  |  |||   |    ||   | | ||     || | |||
Db   7264 GAGTTTCGGTATTGCGAGAAGGTTGTTTCGGTACAGGAGGATCAGTACAAGCTATATGAG 7323

Qy    478 gttagcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaattt 537
          ||||  |||  ||||||    || | |      |||| |||| |  |  |
Db   7324 GTTAAGATTTCAATCTTTTTGAAGTATCTCAGTATAATTTGATATCTATAGAATGGCAGC 7383

Qy    538 tgggttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaacttt 597
          |   | |      |    | |     | |  |  | |    | |||| | | | |
Db   7384 TTTATGATTAAACCCGATGTAAACTGGACATTGAGCAAGCAGTGGTACTGCCTCCATCCC 7443

Qy    598 gattcggattttccaaagtccatgtggataatacgaca---------------------- 635
            || |||   ||| || ||||| ||||||  |   |||
Db   7444 CGTTTGGAAGATCCGAATTCCATATGGATACTGGAACATGTATATATACCGAAGTCAATG 7503

Qy    636 ----ccgctgcacaagccacatatctcctggatccttagagattct-------------- 677
              |||   | || | ||   | || ||  ||| |   |  |  |
Db   7504 CCAGCCGTACCTCAGGGCAAGCACCTACTATATCATCTAATCTCATCTCATCTAATTCTC 7563

Qy    678 ----------------aattataacaatgatatgctttgtcctaacaatcgcagtatatc 721
                          | |||  | ||| |  | | | | || || || ||||| ||||
Db   7564 TGATTTTTCCAGTCGGTACTATTTCCATGGTTGGTTGTATTCTCACGATTGCAGTGTATC 7623

Qy    722 tatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcga 781
          | ||||| ||||| |||||||||| |   ||| || ||  |  | |||| | || |||  |
Db   7624 TCTACATAAAGAAACTGCGAAATTTGCTTGGAAAATGTTTCATTTGCTATGTATTTTGTA 7683

Qy    782 ggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgct 841
           |||| | |||| | | |||  |  |  ||   |||| ||  | | ||||  |||||||
Db   7684 AGTTTGTGCAGTACCTTATCTGGGCAGGAGGAGATTTGAACTTGTGGAATAATATTTGCT 7743

Qy    842 ctccag---------------------------------------------------- 847
          | | ||
Db   7744 CCCTAGCAGGTTAGTTGAATCCATATTAAAGTTTGACTTAAAATATGCGTTTATTCTAAA 7803

Qy    848 ctggttacagctcgcacttttttccggatggcttccaacctctggctctccgtcatcagct 907
          | |||||||      | || ||     |||||||  || | ||||||| || || ||
Db   7804 CAGGTTACACTAACTATTTCTTTGCTTTGGCTTCTCACTTTTGGCTCTCTGTTATGAGTC 7863

Qy    908 accatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaactatcggttcctgc 967
          | || |  ||||||  ||||| ||   |||||||  |  || | || |||||  || ||
Db   7864 ATCAAATATGGAAAAACTTGAGGTTAATCAATCGGGATGAACGTAGCTATCACTTTCTCA 7923

Qy    968 ggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttatatagtta 1027
            ||||||  || ||  ||| ||||| | || ||||||||||| || ||  ||| | ||
Db   7924 TCTACAACATCTACGGTTGGGGCACACCTGCTATCATGACGGCAATTACATATCTGGTGG 7983

Qy   1028 atcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggttttattc 1087
          ||  |  |||||  |    ||   |||  ||| ||| | | ||    || || ||   ||
Db   7984 ATTGGGCCTGGGAGGATAGACCTGACAAACTGAATTGGATTCCCGGTGTTGGCTTATATC 8043

Qy   1088 ggtgctcggtcaaagattgg 1107
          | |||| | | ||   | ||
Db   8044 GATGCTGGATAAACAGTGGG 8063
```

RESULT 8
AC013100
LOCUS AC013100 25416 bp DNA HTG 03-NOV-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION AC013100
VERSION AC013100.1 GI:6223231
KEYWORDS HTG; HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 25416)
AUTHORS Adams,M. and Venter,J.C.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
COMMENT This sequence was identified as CDM:10213216 by the submitter.
For further information on this sequence you may e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

```
FEATURES             Location/Qualifiers
     source          1. .25416
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
BASE COUNT     7806 a   5091 c   4977 g   7542 t
ORIGIN
```

Query Match 10.9%; Score 167.6; DB 63; Length 25416;
Best Local Similarity 50.5%; Pred. No. 1.4e-35;
Matches 871; Conservative 0; Mismatches 629; Indels 226; Gaps 10;

```
Qy     21 tttctgtacagttctgcttttaatttttacgaacaactcaaatgcagatattcccgggtg 80
          ||||||||| || |  |  |||   |  |||| ||||| | |||  | |||||    ||
Db   6016 TTTCTGTACACTTATTATCGTAACCATCGCGAAAAACTCGAGTGCCAAAATTCCGCACTG 6075

Qy     81 caactactacgacacggttgatatctcatacattgaaaggcaaaacgattcgtatttata 140
          ||| ||| | || ||  |  |||| ||  || || || ||| ||||||| | ||| | ||
Db   6076 CAAATACGATGAAACCATCAATATATCCGCACTTTAAACGGCTAAACGATGCCTATATCTA 6135

Qy    141 tgatgacatcgaaattcctgctagcttaactggatactacgagttcaggcagtttgggga 200
          |||  |  | ||||||||||| | | | || ||  | | ||| | ||   || |  |||
Db   6136 TGAACATTTTGAAATTCCTGCCAACCTGACCGGCGAGTTCGACTACAAAGAGCTAATGGA 6195

Qy    201 cggttcgattacgccgattgaa---aagcatttaagggcctgtgtttgcagcgtgcggcc 257
           ||||| |    |||| ||||     | ||||||| | |||| | |||  ||  ||||
Db   6196 TGGTTCCAAAGTTCCGACTGAATTTCCGAATTTAAGAGGCTGTATCTGCAAAGTTAGGCC 6255

Qy    258 ctgtattcgaatctgctgcccagccaagaacttttggccaacggaaaatgcgatgatgg 317
          ||| || ||||| |||||| |   ||| ||  ||||   ||| ||| | ||   ||||||
Db   6256 CTGCATCCGAATTTGCTGCGCCCGCAAAAATATTTTAAGCAATGGAGAGTGTAGTGATGG 6315

Qy    318 tctcaaagagga--------------gctcgcccggttcaaaccctatatatacttcac 362
            | ||  | ||               | | | ||     |  |   ||||||   | ||
Db   6316 CGTTAAGAATGAAATTAAACTAACCATGTTGGACCTAACGATGCAGGATATATTGCTGAC 6375

Qy    363 atacatggacctacaggcacgagtaccact--------------caccgatatggctat 407
            |         ||  || || |       |              ||| ||  ||   ||
Db   6376 GGATCCAACTTTAGCGGAACTAAATATGATACCGCAGTATAACTCCACAGAACTGTTGAT 6435

Qy    408 tatcagagatgagttctttgattgtgatgaaatgatttacatcagcgactttaactattt 467
          | | || ||  ||||       || || || ||  ||  | | |        | | ||
Db   6436 TTTAAGGGAACAGTTTCAGCCATGCGACGAGATTGTTAGCCTAAAAAGAGATGAGTACAC 6495
```

```
Qy    468 cttagaagaggttagcattcaaatcttcaataaatgtggacttatagtt---------tg 518
Db   6496 GATCCTAAAGGTTTGATTCCGTATTTCCTTTGGCTACCCTATCTCATTTGTATCAAACTC 6555

Qy    519 gttccaggatggaaaatttgggttaccgttgacctcttcatggaaaaacaggactactg 578
Db   6556 TTTTAAGGACGGATCCATACTCCTCCACACCTCCGCCGAGATCTTGTCCAATGATCAATA 6615

Qy    579 cttgtaccggcacaactttgattcggattttccaaagtccatgtggataatacgacaccg 638
Db   6616 CTGTCTTTATCCCGAAATCTACTCGGACTTTCCGGAAACCATTAGGATTATAAATCGAAG 6675

Qy    639 ctgcacaagccacatatctcctggatccttagag------------------------- 672
Db   6676 ATGCTACAGAAATGTGATGCCAGGAATCGCCCAGCGTAACAGCACTTGTACAAATCTATT 6735

Qy    673 ------------------------------attctaattataacaatgatatgctttgt 701
Db   6736 ATTCCGATAAGCTTTAATCAATTTTCTTTCAGTTTCTGTGATATCAGTAGTAGGCTTTAT 6795

Qy    702 cctaacaatcgcagtatatctatacattaagaagctgcgaaatgttactggcaagtgcat 761
Db   6796 CCTTACTCTGGCAGTGTATCTTTCTGTTGAGAAGCTTCGGAATTTACTTGGAAAGTGCCT 6855

Qy    762 tgtatgctgtatagtttcgaggtttatccagtgcttgatcatgatactagatcatttaaa 821
Db   6856 TATTTGCTCACTTTTTAGCATGTTCATGGAGTATTTTATTTGGACAATGGATTACTTTCG 6915

Qy    822 tctattgaatggcatttgctctccagctggt---------------------------- 852
Db   6916 TTTATTACAGAGCATTTGTTCCGCTGCAGGTTGGTCTAAAACTAATATTTTAAATTGTAA 6975

Qy    853 ----------------------tacagctcgcacttttttccggatggcttccaacctct 889
Db   6976 TTAAAAATGTATAGCCAATAGGATACATGAAATACTTCTTTTCGATGTCATCGTATCTAT 7035

Qy    890 ggctctccgtcatcagctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgacc 949
Db   7036 GGTTTTCTGTGGTAAGCTTCCATTTGTGGGAACTTTTCACTTCCCTAAATCGTCACGAAC 7095

Qy    950 ctaactatcggttcctgcggtacaacgccttcgtctggagcacagccgcaatcatgacgg 1009
Db   7096 CACAGTATCGATTCCTTATATACAACACCTTTGTCTGGTGCACGGCTGCTATTCCCACAG 7155

Qy   1010 gaagtatttatatagttaatcagatttgggaaaacgatcccagtaaatggaactggttgc 1069
Db   7156 TAGTGATATTTTCAATGAACCAAATGTGGGAAAATGATCCAGGGAAATCTGAATGGTTGC 7215

Qy   1070 ctctggtcggttttattcggtgctcggtcaaagat------------------------ 1104
Db   7216 CTTTGGTCGGTTACTTTGGGTGCTCGGTTAAAGGTGTGTTTTTCATTATAATCGTTCATT 7275

Qy   1105 ------------------------------tggcacccatccgtctggatctatataa 1132
Db   7276 TAAATAGCTTACATATATAATTATTTACAGACTGGAATTCGTCTTCATGGTTCTATAGTC 7335

Qy   1133 gtggaccgtcgctggccctgagcactttcaatgtcgccatgttcgccctgacagccattt 1192
Db   7336 ATATACCAATAGTAATCCTAAACAGCTTTAATGTGATAATGTTCGTCCTAACAGCCATTT 7395

Qy   1193 acattaggaaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggatta 1252
Db   7396 ACATTTGGAAAGTAAAGAAGGGAGTCAAGAGTTTTGCCCAACATGACGAAAGAAATACGA 7455

Qy   1253 actgcataaactttgacagccagacgtaagatatgataaaaccttggaccttggtaaact 1312
Db   7456 CCTGTCTTGAATTCAATGTTCAGACGTAAGAACTAGT------------CTCATTTGAAA 7503

Qy   1313 agttatccttgccgtattatccacatagttacctacagttcctgcggctctccatcgtga 1372
Db   7504 TTATAACTATTTTATATCTTAATTTTAGTTACATACAATTCGTGCGACTTTTTCTAATTA 7563

Qy   1373 tgggccttacttggatattcaatgtcattccgtattctgcacgg----ctccacattttc 1428
Db   7564 TGGGCGCAAGTTGGCTACTTGACCAATTAACGCGATTGGCTGAGGATTCTCACCTTTTGC 7623

Qy   1429 tgggaa--tgggtcgggataatatccgagtattttcacagcgcgtttggaattgtttttgt 1486
Db   7624 TGGACACTATAGTTCTGAACTTAACAGTCTACTTAAATGCCGCCTTTGGAATACTCATAT 7683

Qy   1487 tcgttctgctcgtcctgaagcgcagcacatggactctaatgatgga 1532
Db   7684 TTGTGTTGCTTATACTGAAGGGCAGCACATTTAAGATGATAATGGA 7729
```

```
RESULT   9
AC007754
LOCUS       AC007754   188967 bp    DNA             INV       24-FEB-2001
DEFINITION  Drosophila melanogaster, chromosome 3R, region 87E-87E, BAC clone
            BACR05K08, complete sequence.
ACCESSION   AC007754 AC008360
VERSION     AC007754.4  GI:12830067
KEYWORDS    HTG.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 188967)
  AUTHORS   Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
            Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
            Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
            Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
            Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
            Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
            Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
            Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
            McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
            Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
            Phouanenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
            Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
            Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
  TITLE     Sequencing of Drosophila chromosome 3R, region 87E-87E
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 188967)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
            Rubin,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-1999) Drosophila Genome Center, Lawrence Berkeley
            Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT     On Feb 24, 2001 this sequence version replaced gi:5706740
            gi:5670551.
            Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory, MS 64-121
            Berkeley, CA 94720
            This sequence was assembled using end sequences from a whole genome
            shotgun and from subclones of this BAC and its neighboring clones.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive Web site (http://www.fruitfly.org/sequence/) or send email
            to bdgp@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .188967
                     /organism="Drosophila melanogaster"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /chromosome="3R"
                     /map="87E-87E"
```

```
                /clone="BACR05K08 (D685)"
                /clone_lib="RPCI-98 (Roswell Park Cancer Institute
                Drosophila melanogaster BAC library, partial EcoRI in
                pBACe3.6)"
BASE COUNT     54438 a  39668 c  39881 g  54980 t
ORIGIN

  Query Match           10.9%;  Score 167.6;  DB 4;  Length 188967;
  Best Local Similarity  50.5%;  Pred. No. 1.9e-35;
  Matches  871;  Conservative    0;  Mismatches  629;  Indels  226;  Gaps   10;

Qy      21 tttctgtacagttctgcttttaatttttacgaacaactcaaatgcagatattcccgggtg 80
           |||||||||| || |  |  |||   |  |||| ||||| | |||  | |||||   ||
Db 173191 TTTCTGTACACTTATTATCGTAACCATCGCGAAAAACTCGAGTGCCAAAATTCCGCACTG 173250

Qy      81 caactactacgacacggttgatatctcatacattgaaaggcaaaacgattcgtatttata 140
           ||| ||| | || ||  |  |||| ||  || || || ||| ||||||| | ||| | ||
Db 173251 CAAATACGATGAAACCATCAATATATCGCACTTTAAACGGCTAAACGATGCCTATATCTA 173310

Qy     141 tgatgacatcgaaattcctgctagcttaactggatactacgagttcaggcagtttgggga 200
           |||  |  | ||||||||||| | | || ||  | | ||| | ||   || |  |||
Db 173311 TGAACATTTTGAAATTCCTGCCAACCTGACCGGCGAGTTCGACTACAAAGAGCTAATGGA 173370

Qy     201 cggttcgattacgccgattgaa---aagcatttaagggcctgtgtttgcagcgtgcggcc 257
           |||||  |     |||| ||||     | ||||||| | |||| | ||||  ||  ||||
Db 173371 TGGTTCCAAAGTTCCGACTGAATTTCCGAATTTAAGAGGCTGTATCTGCAAAGTTAGGCC 173430

Qy     258 ctgtattcgaatctgctgcccagccaagaacttttggccaacggaaaatgcgatgatgg 317
           ||| || ||||| |||||| |   ||| ||  ||||   ||| ||| | ||   ||||||
Db 173431 CTGCATCCGAATTTGCTGCGCCCGCAAAAATATTTTAAGCAATGGAGAGTGTAGTGATGG 173490

Qy     318 tctcaaagagga----------------gctcgcccggttcaaaccctatatatacttcac 362
           | ||  | ||                | | ||     |  |   ||||||   | ||
Db 173491 CGTTAAGAATGAAATTAAACTAACCATGTTGGACCTAACGATGCAGGATATATTGCTGAC 173550

Qy     363 atacatggacctacaggcacgagtaccact---------------caccgatatggctat 407
           |        ||  || || |        |                  ||| ||  ||   ||
Db 173551 GGATCCAACTTTAGCGGAACTAAATATGATACCGCAGTATAACTCCACAGAACTGTTGAT 173610

Qy     408 tatcagagatgagttctttgattgtgatgaaatgatttacatcagcgactttaactattt 467
           | | || ||  ||||       || || || ||  ||  | | |       | | ||
Db 173611 TTTAAGGGAACAGTTTCAGCCATGCGACGAGATTGTTAGCCTAAAAAGAGATGAGTACAC 173670

Qy     468 cttagaagaggttagcattcaaatcttcaataaatgtggacttatagtt---------tg 518
           |   | ||||| |  | | |  |    |       |   | ||         |
Db 173671 GATCCTAAAGGTTTGATTCCGTATTTCCTTTGGCTACCCTATCTCATTTGTATCAAACTC 173730

Qy     519 gttccaggatggaaaattttgggttaccgttgacctcttcatggaaaaacaggactactg 578
           ||  |||| |||    |    |   |     | |   ||         | ||  | |
Db 173731 TTTTAAGGACGGATCCATACTCCTCCACACCTCCGCCGAGATCTTGTCCAATGATCAATA 173790

Qy     579 cttgtaccggcacaactttgattcggattttccaaagtccatgtggataatacgacaccg 638
           ||        | | |  |  | ||||| |||||  |  ||||  |||| |||   |   |
Db 173791 CTGTCTTTATCCCGAAATCTACTCGGACTTTCCGGAAACCATTAGGATTATAAATCGAAG 173850

Qy     639 ctgcacaagccacatatctcctggatccttagag-------------------------- 672
           |||   ||  |  |    || |||  |    ||
Db 173851 ATGCTACAGAAATGTGATGCCAGGAATCGCCCAGCGTAACAGCACTTGTACAAATCTATT 173910

Qy     673 ------------------------------attctaattataacaatgatatgctttgt 701
                                           ||    | ||| || |  || |||||
Db 173911 ATTCCGATAAGCTTTAATCAATTTTCTTTCAGTTTCTGTGATATCAGTAGTAGGCTTTAT 173970

Qy     702 cctaacaatcgcagtatatctatacattaagaagctgcgaaatgttactggcaagtgcat 761
           |||  ||  | |||||  ||||| |    || ||||||| || ||| |   ||| ||||||
Db 173971 CCTTACTCTGGCAGTGTATCTTTCTGTTGAGAAGCTTCGGAATTTACTTGGAAAGTGCCT 174030

Qy     762 tgtatgctgtatagtttcgaggtttatccagtgcttgatcatgatactagatcatttaaa 821
           | | ||||   |  ||   | ||| ||  |||  || ||   || | | ||| | ||
Db 174031 TATTTGCTCACTTTTTAGCATGTTCATGGAGTATTTTATTTGGACAATGGATTACTTTCG 174090

Qy     822 tctattgaatggcatttgctctccagctggt---------------------------- 852
```

```
           | ||||  |  |||||| ||  | || |||
Db 174091 TTTATTACAGAGCATTTGTTCCGCTGCAGGTTGGTCTAAAACTAATATTTTAAATTGTAA 174150

Qy     853 ----------------------tacagctcgcacttttttccggatggcttccaacctct 889
                                   ||||      |||| ||   |||| | ||  | || |
Db 174151 TTAAAAATGTATAGCCAATAGGATACATGAAATACTTCTTTTCGATGTCATCGTATCTAT 174210

Qy     890 ggctctccgtcatcagctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgacc 949
           || | || ||  | |||| ||||  |||| || | || || || || ||||   ||| |
Db 174211 GGTTTTCTGTGGTAAGCTTCCATTTGTGGGAACTTTTCACTTCCCTAAATCGTCACGAAC 174270

Qy     950 ctaactatcggttcctgcggtacaacgccttcgtctggagcacagccgcaatcatgacgg 1009
           |  | ||||| |||||    |||||| |||| |||||| |||| || || ||    || |
Db 174271 CACAGTATCGATTCCTTATATACAACACCTTTGTCTGGTGCACGGCTGCTATTCCCACAG 174330

Qy    1010 gaagtatttatatagttaatcagatttgggaaaacgatcccagtaaatggaactggttgc 1069
           |   || | |  | | || || || |||||||| |||||  | ||||   | |||||||
Db 174331 TAGTGATATTTTCAATGAACCAAATGTGGGAAAATGATCCAGGGAAATCTGAATGGTTGC 174390

Qy    1070 ctctggtcggttttattcggtgctcggtcaaagat---------------------- 1104
           || |||||||||   || ||||||||||| |||| |
Db 174391 CTTTGGTCGGTTACTTTGGGTGCTCGGTTAAAGGTGTGTTTTTCATTATAATCGTTCATT 174450

Qy    1105 --------------------------------tggcacccatccgtctggatctatataa 1132
                                            ||| |  | ||   ||| ||||||
Db 174451 TAAATAGCTTACATATATAATTATTTACAGACTGGAATTCGTCTTCATGGTTCTATAGTC 174510

Qy    1133 gtggaccgtcgctggccctgagcactttcaatgtcgccatgttcgccctgacagccattt 1192
           |  |||      |   ||| | ||  || |||||    ||||||| ||| ||||||||||
Db 174511 ATATACCAATAGTAATCCTAAACAGCTTTAATGTGATAATGTTCGTCCTAACAGCCATTT 174570

Qy    1193 acattaggaaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggatta 1252
           ||||| ||||||| ||| |||  | || |  ||  |  |  | || ||  |||  |  |
Db 174571 ACATTTGGAAAGTAAAGAAGGGAGTCAAGAGTTTTGCCCAACATGACGAAAGAAATACGA 174630

Qy    1253 actgcataaactttgacagccagacgtaagatatgataaaaccttggaccttggtaaact 1312
           |||  |  | ||  |   ||||||||||  |  |              ||  |  |
Db 174631 CCTGTCTTGAATTCAATGTTCAGACGTAAGAACTAGT------------CTCATTTGAAA 174678

Qy    1313 agttatccttgccgtattatccacatagttacctacagttcctgcggctctccatcgtga 1372
           || |  |   |||  |     ||||||| |||| ||| |||| || |   |  | |
Db 174679 TTATAACTATTTTATATCTTAATTTTAGTTACATACAATTCGTGCGACTTTTTCTAATTA 174738

Qy    1373 tgggccttacttggatattcaatgtcattccgtattctgcacgg----ctccacattttc 1428
           |||||  | |||| || |  |     |  ||   |  ||   |   |||  | ||| |
Db 174739 TGGGCGCAAGTTGGCTACTTGACCAATTAACGCGATTGGCTGAGGATTCTCACCTTTTGC 174798

Qy    1429 tgggaa--tgggtcgggataatatccgagtattttcacagcgcgtttggaattgttttgt 1486
           |||  |     ||   ||  | | ||  || ||   |  ||| |||||||  |  | ||
Db 174799 TGGACACTATAGTTCTGAACTTAACAGTCTACTTAAATGCCGCCTTTGGAATACTCATAT 174858

Qy    1487 tcgttctgctcgtcctgaagcgcagcacatggactctaatgatgga 1532
           | ||  ||||  | |||||| |||||||||  |   | || |||||
Db 174859 TTGTGTTGCTTATACTGAAGGGCAGCACATTTAAGATGATAATGGA 174904


RESULT  10
AE003699
LOCUS       AE003699   225668 bp    DNA             INV       05-OCT-2000
DEFINITION  Drosophila melanogaster genomic scaffold 142000013386035 section 24
            of 105, complete sequence.
ACCESSION   AE003699 AE002708
VERSION     AE003699.2  GI:10726505
KEYWORDS    HTG.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 225668)
  AUTHORS   Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
            Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
            George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
```

```
            Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
            Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
            Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
            Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
            Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
            Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
            Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
            Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
            Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
            Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
            Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
            Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
            Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S.,
            Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S.,
            Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
            Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
            Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
            Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
            Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
            Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
            Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
            McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
            Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
            Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
            Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
            Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
            Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
            Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
            Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,
            Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
            Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,
            Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
            Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
            Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
            Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
  TITLE     The genome sequence of Drosophila melanogaster
  JOURNAL   Science 287 (5461), 2185-2195 (2000)
  MEDLINE   20196006
REFERENCE   2  (bases 1 to 225668)
  AUTHORS   Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     On Oct 9, 2000 this sequence version replaced gi:7299726.
FEATURES             Location/Qualifiers
     source          1. .225668
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="3R"
     mRNA            join(<14264. .14296,14382. .14587,14883. .14971,15275. .15478,
                     15629. .>17028)
                     /gene="timeout"
                     /product="CT24346"
                     /db_xref="FLYBASE:FBan0008148"
                     /db_xref="FLYBASE:FBgn0038118"
                     /evidence=not_experimental
     gene            <14264. .>17028
                     /gene="timeout"
                     /note="CG8148"
                     /map="87D13-87E1"
                     /db_xref="FLYBASE:FBan0008148"
                     /db_xref="FLYBASE:FBgn0038118"
                     /evidence=not_experimental
     CDS             join(14264. .14296,14382. .14587,14883. .14971,15275. .15478,
                     15629. .17028)
                     /gene="timeout"
                     /note="timeout gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0008148"
                     /db_xref="FLYBASE:FBgn0038118"
                     /evidence=not_experimental
                     /protein_id="AAF54908.1"
                     /db_xref="GI:7299727"
                     /translation="MVEKRGEVAAVAKLFRIFQQVFSVERDVHQEELRRLAIFVVRKF
                     VEVAPTNPKIYAELLFYKGIREANELESGYCDAYEAGTKGAWTEEQESELRFLFEENQ
                     RNPETDKDVIDWILDNLVDKTRNRRTVLKKLKELGLLFKAPTKRSTKSAQSGKNVWQP
                     EEDDELRSLYDQHRIEPDCLERLVNEFAERRSKQQIIKRMLQIHLIADKSEILPAKKG
                     RGKDKPKKDVEMEGEGDEFDFAEEPMFEDQGYKKPKSKPKKVQKRQMVRTPLDVGTIR
                     ALIGQVDSEKYQSAIEWLQECLQDASEDTEEAVEDDDGVPLLPLMENQKNAMEDGDFQ
                     KVLVALGMQPPISGMEAYWRIPIYLNSADLILRSKILAGEEVDAEPEDEAADDEDGEE
                     AEESDEEEDFLEKHSRQRKENIATQQQRKLDSLMFNQSDDETEQRAPPKAKDKPKASE
                     KKAKRKVKGKAKESKDCNDDGTDKAELEKPNTDDLFNQLRAKRATKIKLSDMALNEST
                     NAAQEADDDAFNFNSEDYRARLLELEDEEQNSKDADENGGDKENITNKSLQRARRVNV
                     IDSDSDNDDDGALPADKTNKRPRSDDDGDSGHMDANAEEDSDEEANFLARKKPAKVDG
                     EPAKRRRLAIIDDDDEDDF"
     mRNA            join(33900. .34875,34932. .>35327)
                     /gene="CG8138"
                     /product="CT24326"
                     /db_xref="FLYBASE:FBan0008138"
                     /db_xref="FLYBASE:FBgn0038122"
     gene            <33900. .>35327
                     /gene="CG8138"
                     /map="87E1-87E2"
                     /db_xref="FLYBASE:FBan0008138"
                     /db_xref="FLYBASE:FBgn0038122"
     CDS             join(34570. .34875,34932. .35327)
                     /gene="CG8138"
                     /note="CG8138 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0008138"
                     /db_xref="FLYBASE:FBgn0038122"
                     /protein_id="AAF54909.1"
                     /db_xref="GI:7299728"
                     /translation="MTDFSIAKFNNYQEYLHSFTTVEDFRYLPFHKTVITLTKLGYRE
                     HRTFYEEDEFLNVKKQVEQLLNPTVSAQIYYSQYFKGTDPALWALAEREHGNVQQEIS
                     TIIFLEISGRSGFSKSGYIDFEASLRNYRFKGPNAVDWKAVFEERKLLKPQPSDIVFY
                     DWKTRKIFANDNDNYTVVAHPEHGLMFAHKGDHKNIPVTTKKNPFSSNVRRSMIRSNL
                     YGFIILYDHHVRKKT"
     mRNA            complement(38889. .>39720)
                     /gene="CG8508"
                     /product="CT24847"
                     /db_xref="FLYBASE:FBan0008508"
                     /db_xref="FLYBASE:FBgn0038123"
     gene            complement(<38889. .>39720)
                     /gene="CG8508"
                     /map="87E2-87E2"
                     /db_xref="FLYBASE:FBan0008508"
                     /db_xref="FLYBASE:FBgn0038123"
     CDS             complement(38889. .39692)
                     /gene="CG8508"
                     /note="CG8508 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0008508"
                     /db_xref="FLYBASE:FBgn0038123"
                     /protein_id="AAF54910.1"
                     /db_xref="GI:7299729"
                     /translation="MSLKDLNLTKEQMIKLQQLLAQEIEDESQQERNELSNDDFIANA
                     LDKVQVYVEAGKPEAFVQLKEPLLYKFKITSNGVATGSSESDKFQRADAKEIARWKSD
                     HVINLDCISDEDYEQLSQIQTDESTGETYMLVPIDINILIDDTEITPDTSTKRITEND
                     AYNATGAGHTIYMHPEQEGHQTGLDRNPQAKHACPAPPEEQTNYDYHMHTFEVPKLSF
                     PFELKFGIDTDRLVGGNVDKSAEDSGKSDHDDDLGVATELQDVLEDLEV"
     mRNA            join(<41814. .41864,41916. .42110,42175. .>42450)
                     /gene="CG14380"
                     /product="CT34027"
                     /db_xref="FLYBASE:FBan0014380"
                     /db_xref="FLYBASE:FBgn0038124"
                     /evidence=not_experimental
     gene            <41814. .>42450
                     /gene="CG14380"
                     /map="87E2-87E2"
                     /db_xref="FLYBASE:FBan0014380"
                     /db_xref="FLYBASE:FBgn0038124"
                     /evidence=not_experimental
     CDS             join(41814. .41864,41916. .42110,42175. .42450)
                     /gene="CG14380"
                     /note="CG14380 gene product"
```

```
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0014380"
                     /db_xref="FLYBASE:FBgn0038124"
                     /evidence=not_experimental
                     /protein_id="AAF54911.1"
                     /db_xref="GI:7299730"
                     /translation="MLLSRKEIVSLLLFLTASSAPQIISVLNASVFGNVLPLHVSYEQ
                     LGHQVVCVGTAYICLIYAMSTWLYHSNYSRTRLHMRLVLVAMILIMILFNCASTYYFL
                     TGVSFSKLGDSMELIPLFSSFAMASKSVSLSMSAISVVILIVIMVLLIVTIIQSHRRH
                     KRDRMRYLESQKQ"
     mRNA            join(<44974. .45027,45133. .45464,45518. .>45542)
                     /gene="CG8141"
                     /product="CT24334"
                     /db_xref="FLYBASE:FBan0008141"
                     /db_xref="FLYBASE:FBgn0038125"
                     /evidence=not_experimental
     gene            <44974. .>45542
                     /gene="CG8141"
                     /map="87E2-87E2"
                     /db_xref="FLYBASE:FBan0008141"
                     /db_xref="FLYBASE:FBgn0038125"
                     /evidence=not_experimental
     CDS             join(44974. .45027,45133. .45464,45518. .45542)
                     /gene="CG8141"
                     /note="CG8141 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0008141"
                     /db_xref="FLYBASE:FBgn0038125"
                     /evidence=not_experimental
                     /protein_id="AAF54912.1"
                     /db_xref="GI:7299731"
                     /translation="MKKFQRMKLPRILVMAFKQQQQENDGEDDLLPGTRLRYRRRNLH
                     LDPYVCNECNQYVRGGVITICGHLFCWTCLWPKLSGTAQPRCPCCQRHLLMYEDIMPF
                     HGEGPNARQEDNNVPAQPGSVPRPTGLWVPSDFP"
     mRNA            complement(join(62382. .63015,63296. .63554,64122. .64322,
                     67319. .>67690))
                     /gene="CG8483"
                     /product="CT24811"
                     /db_xref="FLYBASE:FBan0008483"

  Query Match             10.9%;  Score 167.6;  DB 5;  Length 225668;
  Best Local Similarity   50.5%;  Pred. No. 1.9e-35;
  Matches  871;  Conservative    0;  Mismatches  629;  Indels  226;  Gaps   10;

Qy      21 tttctgtacagttctgcttttaatttttacgaacaactcaaatgcagatattcccgggtg 80
           |||||||||| || |  |  |||   |  |||| ||||| | |||  | |||||   ||
Db 207021 TTTCTGTACACTTATTATCGTAACCATCGCGAAAAACTCGAGTGCCAAAATTCCGCACTG 207080

Qy      81 caactactacgacacggttgatatctcatacattgaaaggcaaaacgattcgtatttata 140
           ||| ||| | || ||  |  |||| ||  || || || ||| ||||||| | ||| | ||
Db 207081 CAAATACGATGAAACCATCAATATATCGCACTTTAAACGGCTAAACGATGCCTATATCTA 207140

Qy     141 tgatgacatcgaaattcctgctagcttaactggatactacgagttcaggcagtttgggga 200
           |||  |  | ||||||||||| | | | || ||  | | ||| | ||   || |   |||
Db 207141 TGAACATTTTGAAATTCCTGCCAACCTGACCGGCGAGTTCGACTACAAAGAGCTAATGGA 207200

Qy     201 cggttcgattacgccgattgaa---aagcatttaagggcctgtgtttgcagcgtgcggcc 257
           |||||  |     |||| ||||     | ||||||| | |||| | |||| ||  ||||
Db 207201 TGGTTCCAAAGTTCCGACTGAATTTCCGAATTTAAGAGGCTGTATCTGCAAAGTTAGGCC 207260

Qy     258 ctgtattcgaatctgctgcccagccaagaacttttggccaacggaaaatgcgatgatgg 317
           | ||| || ||||| |||||| |   ||| ||  ||||   ||| ||| | ||  ||||||
Db 207261 CTGCATCCGAATTTGCTGCGCCCGCAAAAATATTTTAAGCAATGGAGAGTGTAGTGATGG 207320

Qy     318 tctcaaagagga----------------gctcgcccggttcaaaccctatatatacttcac 362
             | ||  | ||                | | | ||     |  |   ||||||   | ||
Db 207321 CGTTAAGAATGAAATTAAACTAACCATGTTGGACCTAACGATGCAGGATATATTGCTGAC 207380

Qy     363 atacatggacctacaggcacgagtaccact----------------caccgatatggctat 407
             |        ||  || || |        |                ||| ||  ||   ||
Db 207381 GGATCCAACTTTAGCGGAACTAAATATGATACCGCAGTATAACTCCACAGAACTGTTGAT 207440

Qy     408 tatcagagatgagttctttgattgtgatgaaatgatttacatcagcgactttaactattt 467
```

```
                 | | || ||  ||||       || || || ||  ||  | | |       | | ||
Db 207441 TTTAAGGGAACAGTTTCAGCCATGCGACGAGATTGTTAGCCTAAAAAGAGATGAGTACAC 207500

Qy     468 cttagaagaggttagcattcaaatcttcaataaatgtggacttatagtt---------tg 518
             |   | ||||| |  | |  || | |  |  |       |   | ||         |
Db 207501 GATCCTAAAGGTTTGATTCCGTATTTCCTTTGGCTACCCTATCTCATTTGTATCAAACTC 207560

Qy     519 gttccaggatggaaaattttgggttaccgttgacctcttcatggaaaaacaggactactg 578
           | ||  |||| |||    |      |   |      |  |   ||         | ||  | |
Db 207561 TTTTAAGGACGGATCCATACTCCTCCACACCTCCGCCGAGATCTTGTCCAATGATCAATA 207620

Qy     579 cttgtaccggcacaactttgattcggattttccaaagtccatgtggataatacgacaccg 638
           ||         | |  |   | ||||| ||||  |  ||||  |||| |||   |   |
Db 207621 CTGTCTTTATCCCGAAATCTACTCGGACTTTCCGGAAACCATTAGGATTATAAATCGAAG 207680

Qy     639 ctgcacaagccacatatctcctggatccttagag-------------------------- 672
           |||   ||  |  |    || |||  |    ||
Db 207681 ATGCTACAGAAATGTGATGCCAGGAATCGCCCAGCGTAACAGCACTTGTACAAATCTATT 207740

Qy     673 ------------------------------attctaattataacaatgatatgctttgt 701
                                           ||    | ||| || |  || ||||| |
Db 207741 ATTCCGATAAGCTTTAATCAATTTTCTTTCAGTTTCTGTGATATCAGTAGTAGGCTTTAT 207800

Qy     702 cctaacaatcgcagtatatctatacattaagaagctgcgaaatgttactggcaagtgcat 761
           ||| ||  | ||||| ||||| |   || |||||||| || ||| |   ||| ||||| |
Db 207801 CCTTACTCTGGCAGTGTATCTTTCTGTTGAGAAGCTTCGGAATTTACTTGGAAAGTGCCT 207860

Qy     762 tgtatgctgtatagtttcgaggtttatccagtgcttgatcatgatactagatcatttaaa 821
           | | ||||   |  ||   | ||| ||  |||  || ||   || | | ||| | ||
Db 207861 TATTTGCTCACTTTTTAGCATGTTCATGGAGTATTTTATTTGGACAATGGATTACTTTCG 207920

Qy     822 tctattgaatggcatttgctctccagctggt---------------------------- 852
           | ||||  |  |||||||  ||  | ||  |||
Db 207921 TTTATTACAGAGCATTTGTTCCGCTGCAGGTTGGTCTAAAACTAATATTTTAAATTGTAA 207980

Qy     853 -----------------------tacagctcgcacttttttccggatggcttccaacctct 889
                                   ||||      |||| ||   |||| | ||  | || |
Db 207981 TTAAAAATGTATAGCCAATAGGATACATGAAATACTTCTTTTCGATGTCATCGTATCTAT 208040

Qy     890 ggctctccgtcatcagctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgacc 949
           || | || ||  | |||| ||||  |||| || | || || || || ||||||   ||| |
Db 208041 GGTTTTCTGTGGTAAGCTTCCATTTGTGGGAACTTTTCACTTCCCTAAATCGTCACGAAC 208100

Qy     950 ctaactatcggttcctgcggtacaacgccttcgtctggagcacagccgcaatcatgacgg 1009
           |  | ||||| |||||    |||||| |||| |||||| |||| || || ||    || |
Db 208101 CACAGTATCGATTCCTTATATACAACACCTTTGTCTGGTGCACGGCTGCTATTCCCACAG 208160

Qy    1010 gaagtatttatatagttaatcagatttgggaaaacgatcccagtaaatggaactggttgc 1069
           |   || | |  | || || || |||||||| |||||  | ||||   | |||||||
Db 208161 TAGTGATATTTTCAATGAACCAAATGTGGGAAAATGATCCAGGGAAATCTGAATGGTTGC 208220

Qy    1070 ctctggtcggttttattcggtgctcggtcaaagat------------------------- 1104
           || |||||||||   || ||||||||||| |||| |
Db 208221 CTTTGGTCGGTTACTTTGGGTGCTCGGTTAAAGGTGTGTTTTTCATTATAATCGTTCATT 208280

Qy    1105 --------------------------------tggcacccatccgtctggatctatataa 1132
                                             ||| |  | ||    ||| ||||||
Db 208281 TAAATAGCTTACATATATAATTATTTACAGACTGGAATTCGTCTTCATGGTTCTATAGTC 208340

Qy    1133 gtggaccgtcgctggccctgagcactttcaatgtcgccatgttcgccctgacagccattt 1192
           |  |||      |   ||| | ||  || |||||    ||||||| ||| ||||||||||
Db 208341 ATATACCAATAGTAATCCTAAACAGCTTTAATGTGATAATGTTCGTCCTAACAGCCATTT 208400

Qy    1193 acattaggaaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggatta 1252
           |||||  ||||||| |||  |||  | || |  ||  |  |  | || ||  |||  |  |
Db 208401 ACATTTGGAAAGTAAAGAAGGGAGTCAAGAGTTTTGCCCAACATGACGAAAGAAATACGA 208460

Qy    1253 actgcataaactttgacagccagacgtaagatatgataaaaccttggaccttggtaaact 1312
           |||  |  | ||  |    ||||||||||  |  |            ||   |  |
Db 208461 CCTGTCTTGAATTCAATGTTCAGACGTAAGAACTAGT-----------CTCATTTGAAA 208508

Qy    1313 agttatccttgccgtattatccacatagttacctacagttcctgcggctctccatcgtga 1372
             || |  |    |||  |     ||||||| |||| ||| |||| || |   |  | |
```

```
Db 208509 TTATAACTATTTTATATCTTAATTTTAGTTACATACAATTCGTGCGACTTTTTCTAATTA 208568

Qy   1373 tgggccttacttggatattcaatgtcattccgtattctgcacgg----ctccacattttc 1428
          |||||   | |||| || |  |      |  ||   |  ||   |    |||  | ||| |
Db 208569 TGGGCGCAAGTTGGCTACTTGACCAATTAACGCGATTGGCTGAGGATTCTCACCTTTTGC 208628

Qy   1429 tgggaa--tgggtcgggataatatccgagtattttcacagcgcgtttggaattgttttgt 1486
          |||  |     ||   ||   || | |  || ||  |   ||| ||||||||  |  | |
Db 208629 TGGACACTATAGTTCTGAACTTAACAGTCTACTTAAATGCCGCCTTTGGAATACTCATAT 208688

Qy   1487 tcgttctgctcgtcctgaagcgcagcacatggactctaatgatgga 1532
          | ||  |||| | |||||| |||||||||  |   | || |||||
Db 208689 TTGTGTTGCTTATACTGAAGGGCAGCACATTTAAGATGATAATGGA 208734


RESULT  11
AC010557
LOCUS       AC010557   153511 bp    DNA              HTG       16-OCT-1999
DEFINITION  Drosophila melanogaster chromosome 3L/65F1 clone RPCI98-9E21, ***
            SEQUENCING IN PROGRESS ***, 58 unordered pieces.
ACCESSION   AC010557
VERSION     AC010557.2  GI:5916433
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 153511)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lu,J., Lucier,R.,
            Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M., Morris,S.,
            Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G.,
            Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L., Quiles,M.,
            Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S., Shah,E.,
            Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R., Tabor,P.,
            Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M., Watlington,S.,
            Weinstock,G., Weinstock,I.R., Williamson,A., Worley,K., Wren,J.,
            Wrensford,G., Yu,W., Zhou,X., Nelson,D. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 153511)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-SEP-1999) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 21, 1999 this sequence version replaced gi:5902223.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 58 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1      846: contig of 846 bp in length
            *     847     1679: contig of 833 bp in length
            *    1680     2509: contig of 830 bp in length
            *    2510     3394: contig of 885 bp in length
            *    3395     4267: contig of 873 bp in length
            *    4268     5541: contig of 1274 bp in length
            *    5542     6396: contig of 855 bp in length
            *    6397     7261: contig of 865 bp in length
            *    7262     8343: contig of 1082 bp in length
            *    8344     9413: contig of 1070 bp in length
            *    9414    10865: contig of 1452 bp in length
            *   10866    12009: contig of 1144 bp in length
            *   12010    13519: contig of 1510 bp in length
            *   13520    14348: contig of 829 bp in length
            *   14349    15937: contig of 1589 bp in length
            *   15938    16887: contig of 950 bp in length
            *   16888    18554: contig of 1667 bp in length
            *   18555    19629: contig of 1075 bp in length
            *   19630    20577: contig of 948 bp in length
            *   20578    22711: contig of 2134 bp in length
            *   22712    24833: contig of 2122 bp in length
            *   24834    27049: contig of 2216 bp in length
            *   27050    28883: contig of 1834 bp in length
            *   28884    30953: contig of 2070 bp in length
            *   30954    32553: contig of 1600 bp in length
            *   32554    34695: contig of 2142 bp in length
            *   34696    36584: contig of 1889 bp in length
            *   36585    39011: contig of 2427 bp in length
            *   39012    41368: contig of 2357 bp in length
            *   41369    43504: contig of 2136 bp in length
            *   43505    46362: contig of 2858 bp in length
            *   46363    48537: contig of 2175 bp in length
            *   48538    51297: contig of 2760 bp in length
            *   51298    54270: contig of 2973 bp in length
            *   54271    57488: contig of 3218 bp in length
            *   57489    61396: contig of 3908 bp in length
            *   61397    63990: contig of 2594 bp in length
            *   63991    66330: contig of 2340 bp in length
            *   66331    69133: contig of 2803 bp in length
            *   69134    72769: contig of 3636 bp in length
            *   72770    75116: contig of 2347 bp in length
            *   75117    79424: contig of 4308 bp in length
            *   79425    83016: contig of 3592 bp in length
            *   83017    86336: contig of 3320 bp in length
            *   86337    89806: contig of 3470 bp in length
            *   89807    93766: contig of 3960 bp in length
            *   93767    96238: contig of 2472 bp in length
            *   96239   101446: contig of 5208 bp in length
            *  101447   105351: contig of 3905 bp in length
            *  105352   108404: contig of 3053 bp in length
            *  108405   112509: contig of 4105 bp in length
            *  112510   117380: contig of 4871 bp in length
            *  117381   123170: contig of 5790 bp in length
            *  123171   128370: contig of 5200 bp in length
            *  128371   133042: contig of 4672 bp in length
            *  133043   139150: contig of 6108 bp in length
            *  139151   143661: contig of 4511 bp in length
            *  143662   153511: contig of 9850 bp in length.
FEATURES             Location/Qualifiers
     source          1. .153511
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="3L/65F1"
                     /clone="RPCI98-9E21"
BASE COUNT    42964 a  32719 c  33190 g  44520 t    118 others
ORIGIN


  Query Match             9.4%;  Score 144.8;  DB 61;  Length 153511;
  Best Local Similarity   53.9%;  Pred. No. 4.3e-29;
  Matches  344;  Conservative    0;  Mismatches  287;  Indels    7;  Gaps    2;

Qy      4 aggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactcaaat 63
          ||  | | |  |  |||||   || | ||||  | ||  |  || | | | | ||  |
Db  56733 AGAATGCTTCTGAATATTTTAGCTATAATTCTTGTATTCGTAATTTCAAGCCAATCGGAG 56792

Qy     64 gcagatattcccgggtgcaactactacgacacggttgatatctcatacattgaaaggcaa 123
          |||| |||||| || ||  | |||| |||||| || ||||| ||  | ||   || || |
Db  56793 GCAGTTATTCCTGGATGTGATTACTTCGACACCGTGGATATATCTCATATCCCAAAGCTA 56852

Qy    124 aacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactacgag 183
          || || || |||   || || ||  |   |||||||||    | || ||    ||
```

```
Db  56853 AATGACTCCTATGCGTACGAGGAGCTAATCATTCCTGCTCATCTCACGGGTCTATATACA 56912

Qy    184 ttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgtgtt 243
          || |||||  | | ||| || ||      |||  |  | |  || ||||| || ||| |
Db  56913 TTTAGGCAACTGGCGGATGGATCACAGGAGCCAGTGAAGAGTCACTTAAGAGCATGTATC 56972

Qy    244 tgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttttggccaacgga 303
          ||||   |    || ||||| ||| ||||||||||   ||| ||  |  || | ||  |
Db  56973 TGCAAACTAAAACCTTGTATCCGATTCTGCTGCCCCCGCAACAAAATGATGCCGAATAGT 57032

Qy    304 aaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttcaca 363
            ||    ||||| |||| |||| | ||    ||| |||| |||||  |  |  |||||
Db  57033 CGCTGTAGCGATGGACTCACAGAGAATCTGAAGCGGATCAATCCCTACTTGAAGATCACA 57092

Qy    364 tacatggacctacaggcacgagta------ccactcaccgatatggctattatcagagat 417
              |||   |     | ||  |       |  |||||||| |||    |  | ||| ||
Db  57093 CTGGAGGATGGAACTATAGGAAAATATTACCTTCTCACCGACATGATCGTCTTAAGATAT 57152

Qy    418 gagttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagag 477
          |||||     |||| ||  |  |  ||     ||| ||   | | ||     || | |||
Db  57153 GAGTTTCGGTATTGCGAGAAGGTTGTTCGGTACAG-GAGGATCAGTACAAGCTATATGAG 57211

Qy    478 gttagcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaattt 537
          ||||  |||  |||||||    || | |     |||| |||| |  |   |
Db  57212 GTTAAGATTTCAATCTTTTTGAAGTATCTCAGTATAATTTGATATCTATAGAATGGCAGC 57271

Qy    538 tgggttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaacttt 597
          |   | |       |    | |     | |  |  | |    | |||| | | |
Db  57272 TTTATGATTAAACCCGATGTAAACTGGACATTGAGCAAGCAGTGGTACTGCCTCCATCCC 57331

Qy    598 gattcggattttccaaagtccatgtggataatacgaca 635
            || |||   ||| || ||||| ||||||  |   |||
Db  57332 CGTTTGGAAGATCCGAATTCCATATGGATACTGGAACA 57369


RESULT  12
AC008360
ID   AC008360   standard; DNA; HTG; 121857 BP.
XX
AC   AC008360;
XX
SV   AC008360.2
XX
DT   03-AUG-1999 (Rel. 60, Created)
DT   03-MAR-2000 (Rel. 62, Last updated, Version 4)
XX
DE   Drosophila melanogaster chromosome 3 clone BACR06H18 (D979) RPCI-98 06.H.18
DE   map 87E-87F strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***, 102
DE   unordered pieces.
XX
KW   HTG; HTGS_PHASE1.
XX
OS   Drosophila melanogaster (fruit fly)
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
OC   Drosophilidae; Drosophila.
XX
RN   [1]
RP   1-121857
RA   Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
RA   Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA   Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA   Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA   Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA   Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA   Sethi H., Snir E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA   Zieran L.L., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster";
RL   Unpublished.
XX
RN   [2]
RP   1-121857
RA   Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
RA   Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA   Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA   Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA   Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA   Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA   Sethi H., Snir E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA   Zieran L.L., Rubin G.M.;
RT   ;
RL   Submitted (02-AUG-1999) to the EMBL/GenBank/DDBJ databases.
RL   Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121,
RL   Berkeley, CA 94720, USA
XX
CC   On Aug 6, 1999 this sequence version replaced gi:5670356.
CC   For further information about this sequence, including its location
CC   and relationship to other sequences, please visit our sequence
CC   archive Web site (http://www.fruitfly.org/sequence/) or send email
CC   to bdgp@fruitfly.berkeley.edu. All contigs in this submission meet
CC   the following cutoffs: length >= 200 bases.
CC   * NOTE: This is a 'working draft' sequence. It currently
CC   * consists of 102 contigs. The true order of the pieces
CC   * is not known and their order in this sequence record is
CC   * arbitrary. Gaps between the contigs are represented as
CC   * runs of N, but the exact sizes of the gaps are unknown.
CC   * This record will be updated with the finished sequence
CC   * as soon as it is available and the accession number will
CC   * be preserved.
CC   *        1      645: contig of 645 bp in length
CC   *      646      725: gap of unknown length
CC   *      726     1138: contig of 413 bp in length
CC   *     1139     1218: gap of unknown length
CC   *     1219     1922: contig of 704 bp in length
CC   *     1923     2002: gap of unknown length
CC   *     2003     2957: contig of 955 bp in length
CC   *     2958     3037: gap of unknown length
CC   *     3038     4306: contig of 1269 bp in length
CC   *     4307     4386: gap of unknown length
CC   *     4387     4983: contig of 597 bp in length
CC   *     4984     5063: gap of unknown length
CC   *     5064     5771: contig of 708 bp in length
CC   *     5772     5851: gap of unknown length
CC   *     5852     6387: contig of 536 bp in length
CC   *     6388     6467: gap of unknown length
CC   *     6468     7235: contig of 768 bp in length
CC   *     7236     7315: gap of unknown length
CC   *     7316     8012: contig of 697 bp in length
CC   *     8013     8092: gap of unknown length
CC   *     8093     9087: contig of 995 bp in length
CC   *     9088     9167: gap of unknown length
CC   *     9168    10351: contig of 1184 bp in length
CC   *    10352    10431: gap of unknown length
CC   *    10432    10984: contig of 553 bp in length
CC   *    10985    11064: gap of unknown length
CC   *    11065    11736: contig of 672 bp in length
CC   *    11737    11816: gap of unknown length
CC   *    11817    12984: contig of 1168 bp in length
CC   *    12985    13064: gap of unknown length
CC   *    13065    13756: contig of 692 bp in length
CC   *    13757    13836: gap of unknown length
CC   *    13837    14777: contig of 941 bp in length
CC   *    14778    14857: gap of unknown length
CC   *    14858    15767: contig of 910 bp in length
CC   *    15768    15847: gap of unknown length
CC   *    15848    16967: contig of 1120 bp in length
CC   *    16968    17047: gap of unknown length
CC   *    17048    17743: contig of 696 bp in length
CC   *    17744    17823: gap of unknown length
CC   *    17824    18439: contig of 616 bp in length
CC   *    18440    18519: gap of unknown length
CC   *    18520    19400: contig of 881 bp in length
CC   *    19401    19480: gap of unknown length
CC   *    19481    20159: contig of 679 bp in length
CC   *    20160    20239: gap of unknown length
CC   *    20240    20980: contig of 741 bp in length
```

```
CC   *    20981    21060: gap of unknown length
CC   *    21061    21966: contig of 906 bp in length
CC   *    21967    22046: gap of unknown length
CC   *    22047    23217: contig of 1171 bp in length
CC   *    23218    23297: gap of unknown length
CC   *    23298    24686: contig of 1389 bp in length
CC   *    24687    24766: gap of unknown length
CC   *    24767    26287: contig of 1521 bp in length
CC   *    26288    26367: gap of unknown length
CC   *    26368    27055: contig of 688 bp in length
CC   *    27056    27135: gap of unknown length
CC   *    27136    28209: contig of 1074 bp in length
CC   *    28210    28289: gap of unknown length
CC   *    28290    29499: contig of 1210 bp in length
CC   *    29500    29579: gap of unknown length
CC   *    29580    30447: contig of 868 bp in length
CC   *    30448    30527: gap of unknown length
CC   *    30528    31993: contig of 1466 bp in length
CC   *    31994    32073: gap of unknown length
CC   *    32074    33251: contig of 1178 bp in length
CC   *    33252    33331: gap of unknown length
CC   *    33332    34659: contig of 1328 bp in length
CC   *    34660    34739: gap of unknown length
CC   *    34740    35636: contig of 897 bp in length
CC   *    35637    35716: gap of unknown length
CC   *    35717    36791: contig of 1075 bp in length
CC   *    36792    36871: gap of unknown length
CC   *    36872    38045: contig of 1174 bp in length
CC   *    38046    38125: gap of unknown length
CC   *    38126    39250: contig of 1125 bp in length
CC   *    39251    39330: gap of unknown length
CC   *    39331    40706: contig of 1376 bp in length
CC   *    40707    40786: gap of unknown length
CC   *    40787    41575: contig of 789 bp in length
CC   *    41576    41655: gap of unknown length
CC   *    41656    42837: contig of 1182 bp in length
CC   *    42838    42917: gap of unknown length
CC   *    42918    44079: contig of 1162 bp in length
CC   *    44080    44159: gap of unknown length
CC   *    44160    45350: contig of 1191 bp in length
CC   *    45351    45430: gap of unknown length
CC   *    45431    46870: contig of 1440 bp in length
CC   *    46871    46950: gap of unknown length
CC   *    46951    48162: contig of 1212 bp in length
CC   *    48163    48242: gap of unknown length
CC   *    48243    50081: contig of 1839 bp in length
CC   *    50082    50161: gap of unknown length
CC   *    50162    52341: contig of 2180 bp in length
CC   *    52342    52421: gap of unknown length
CC   *    52422    53632: contig of 1211 bp in length
CC   *    53633    53712: gap of unknown length
CC   *    53713    55481: contig of 1769 bp in length
CC   *    55482    55561: gap of unknown length
CC   *    55562    56444: contig of 883 bp in length
CC   *    56445    56524: gap of unknown length
CC   *    56525    57726: contig of 1202 bp in length
CC   *    57727    57806: gap of unknown length
CC   *    57807    58977: contig of 1171 bp in length
CC   *    58978    59057: gap of unknown length
CC   *    59058    60477: contig of 1420 bp in length
CC   *    60478    60557: gap of unknown length
CC   *    60558    62129: contig of 1572 bp in length
CC   *    62130    62209: gap of unknown length
CC   *    62210    63549: contig of 1340 bp in length
CC   *    63550    63629: gap of unknown length
CC   *    63630    65406: contig of 1777 bp in length
CC   *    65407    65486: gap of unknown length
CC   *    65487    66262: contig of 776 bp in length
CC   *    66263    66342: gap of unknown length
CC   *    66343    67434: contig of 1092 bp in length
CC   *    67435    67514: gap of unknown length
CC   *    67515    68875: contig of 1361 bp in length
CC   *    68876    68955: gap of unknown length
CC   *    68956    70478: contig of 1523 bp in length
CC   *    70479    70558: gap of unknown length
CC   *    70559    71293: contig of 735 bp in length
CC   *    71294    71373: gap of unknown length
CC   *    71374    73982: contig of 2609 bp in length
CC   *    73983    74062: gap of unknown length
CC   *    74063    75285: contig of 1223 bp in length
CC   *    75286    75365: gap of unknown length
CC   *    75366    78399: contig of 3034 bp in length
CC   *    78400    78479: gap of unknown length
CC   *    78480    80398: contig of 1919 bp in length
CC   *    80399    80478: gap of unknown length
CC   *    80479    83078: contig of 2600 bp in length
CC   *    83079    83158: gap of unknown length
CC   *    83159    86699: contig of 3541 bp in length
CC   *    86700    86779: gap of unknown length
CC   *    86780    91120: contig of 4341 bp in length
CC   *    91121    91200: gap of unknown length

  Query Match              8.4%;  Score 129.4;  DB 30;  Length 121857;
  Best Local Similarity   58.5%;  Pred. No. 8.3e-25;
  Matches  279;  Conservative    0;  Mismatches  146;  Indels   52;  Gaps    1;

Qy     680 ttataacaatgatatgctttgtcctaacaatcgcagtatatctatacattaagaagctgc 739
           | ||| || |  || ||||| |||| ||  | ||||| ||||| |   || ||||||| |
Db   12391 TGATATCAGTAGTAGGCTTTATCCTTACTCTGGCAGTGTATCTTTCTGTTGAGAAGCTTC 12450

Qy     740 gaaatgttactggcaagtgcattgtatgctgtatagtttcgaggtttatccagtgcttga 799
           | ||| |   ||| |||||| || | ||||   |  ||   | ||| ||  |||  || |
Db   12451 GGAATTTACTTGGAAAGTGCCTTATTTGCTCACTTTTTAGCATGTTCATGGAGTATTTTA 12510

Qy     800 tcatgatactagatcatttaaatctattgaatggcatttgctctccagctggt------- 852
           |   || | | ||| | ||   | ||||  |  ||||||| ||  | || |||
Db   12511 TTTGGACAATGGATTACTTTCGTTTATTACAGAGCATTTGTTCCGCTGCAGGTTGGTCTA 12570

Qy     853 ----------------------------------------------tacagctcgcacttt 867
                                                          ||||   ||||
Db   12571 AAACTAATATTTTAAATTGTAATTAAAAATGTATAGCCAATAGGATACATGAAATACTTC 12630

Qy     868 ttccggatggcttccaacctctggctctccgtcatcagctaccatacgtggaaagtcttg 927
           ||   |||| | ||  | || ||| | || ||  | |||| ||||  |||| || | ||
Db   12631 TTTTCGATGTCATCGTATCTATGGTTTTCTGTGGTAAGCTTCCATTTGTGGGAACTTTTC 12690

Qy     928 acgtcgctcaatcgagtcgaccctaactatcggttcctgcggtacaacgccttcgtctgg 987
           || || || |||||   ||| ||  | ||||| |||||    |||||| |||| ||||||
Db   12691 ACTTCCCTAAATCGTCACGAACCACAGTATCGATTCCTTATATACAACACCTTTGTCTGG 12750

Qy     988 agcacagccgcaatcatgacgggaagtatttatatagttaatcagatttgggaaaacgat 1047
           ||||  || || ||    || | |  || | |  | | || || || |||||||| |||
Db   12751 TGCACGGCTGCTATTCCCACAGTAGTGATATTTTCAATGAACCAAATGTGGGAAAATGAT 12810

Qy    1048 cccagtaaatggaactggttgcctctggtcggttttattcggtgctcggtcaaagat 1104
           ||  | ||||   | |||||||| |||||||||   || ||||||||| |||| |
Db   12811 CCAGGGAAATCTGAATGGTTGCCTTTGGTCGGTTACTTTGGGTGCTCGGTTAAAGGT 12867


RESULT  13
AF109308
LOCUS       AF109308     1948 bp    mRNA            INV       03-MAR-1999
DEFINITION  Drosophila melanogaster putative G-protein-coupled receptor (mth)
            mRNA, complete cds.
ACCESSION   AF109308
VERSION     AF109308.1  GI:4324671
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1948)
  AUTHORS   Lin,Y.J., Seroude,L. and Benzer,S.
  TITLE     Extended life-span and stress resistance in the Drosophila mutant
            methuselah
```

```
  JOURNAL   Science 282 (5390), 943-946 (1998)
  MEDLINE   99011353
REFERENCE   2  (bases 1 to 1948)
  AUTHORS   Lin,Y.J., Seroude,L. and Benzer,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-NOV-1998) Biology, California Institute of
            Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA
FEATURES             Location/Qualifiers
     source          1. .1948
                     /organism="Drosophila melanogaster"
                     /strain="Canton S"
                     /db_xref="taxon:7227"
                     /chromosome="3"
                     /map="61C"
                     /clone="LD08316"
     gene            1. .1948
                     /gene="mth"
                     /note="methuselah; a P-element insertion into the
                     methuselah gene results in extended life-span"
     CDS             109. .1653
                     /gene="mth"
                     /codon_start=1
                     /product="putative G-protein-coupled receptor"
                     /protein_id="AAD16981.1"
                     /db_xref="GI:4324672"
                     /translation="MKTLLVLRISTVILVVLVIQKSYADILECDYFDTVDISAAQKLQ
                     NGSYLFEGLLVPAILTGEYDFRILPDDSKQKVARHIRGCVCKLKPCVRFCCPHDHIMD
                     NGVCYDNMSDEELAELDPFLNVTLDDGSVSRRHFKNELIVQWDLPMPCDGMFYLDNRE
                     EQDKYTLFENGTFFRHFDRVTLRKREYCLQHLTFADGNATSIRIAPHNCLIVPSITGQ
                     TVVMISSLICMVLTIAVYLFVKKLQNLHGKCFICYMVCLFMGYLFLLLDLWQISISFC
                     KPAGFLGYFFVMAAFFWLSVISLHLWNTFRGSSHKANRFLFEHRFLAYNTYAWGMAVV
                     LTGITVLADNIVENQDWNPRVGHEGHCWIYTQAWSAMLYFYGPMVFLIAFNITMFILT
                     AKRILGVKKDIQNFAHRQERKQKLNSDKQTYTFFLRLFIIMGLSWSLEIGSYFSQSNQ
                     TWANVFLVADYLNWSQGIIIFILFVLKRSTWRLLQESIRGEGEEVNNSEEEISLENTT
                     TRNVLL"
BASE COUNT      550 a    421 c    420 g    557 t
ORIGIN
```

```
  Query Match             8.1%;  Score 124.8;  DB 5;  Length 1948;
  Best Local Similarity   49.0%;  Pred. No. 8.6e-24;
  Matches  611;  Conservative    0;  Mismatches  572;  Indels   63;  Gaps    8;

Qy      32 ttctgcttttaatttttacgaacaactcaaatgcagatattcccgggtgcaactactacg 91
           | ||  || |  |  |||   | || ||| ||||||| |||| ||  ||| | || | ||
Db     146 TACTTGTTGTGCTGGTTATTCAAAAATCATATGCAGACATTCTCGAATGCGATTATTTCG 205

Qy      92 acacggttgatatctcatacattgaaaggcaaaacgat---tcgtatttatatgatgaca 148
           |||| || ||||| ||      | |||  | | |  ||   ||||| |||| ||| |
Db     206 ACACTGTCGATATTTCGGCAGCTCAAAAACTACAGAATGGATCGTACTTATTTGAGGGGT 265

Qy     149 tcgaaattcctgctagcttaactggatactacgagttcaggcagtttggggacggttcga 208
           |     |||| || |  || || ||| | || || || |||    |  |||||  ||||
Db     266 TGCTCGTTCCGGCCATTTTGACGGGAGAATATGACTTTAGGATTCTCCCGGACGACTCGA 325

Qy     209 ttacgccgattgaaaagcatttaagggcctgtgtttgcagcgtgcggccctgtattcgaa 268
           |  | | |     ||  |||| |  ||||| |||   ||  |||||||  |  ||
Db     326 AGCAGAAGGTGGCTCGTCACATAAGAGGATGTGTGTGCAAGCTGAAGCCCTGTGTCAGAT 385

Qy     269 tctgctgcccagccaagaacttttggccaacggaaaatgcgatgatggtct---caaag 325
           | |||||||   | |  |  |  |||  || ||    ||| | ||     |   | | |
Db     386 TTTGCTGCCCTCACGACCATATCATGGATAATGGTGTCTGCTACGACAACATGTCCGACG 445

Qy     326 aggagctcgcccggttcaaaccctatatatacttcacatacatggacctacaggcacgag 385
           |||||||  ||     ||  | ||||   |  |  | ||       |||     ||
Db     446 AGGAGCTGGCTGAACTCGATCCCTTCCTTAATGTGACTCTTGACGACGGGTCGGTGTCCC 505

Qy     386 taccactcaccgatatggctattatcagagatgagttctttgattgtgatgaaatgattt 445
           |  ||    | |  |   | |||     |   |
Db     506 GCAGACACTTTAAAAACGAACTAATCGTCCAGTGG------------------------ 540

Qy     446 acatcagcgactttaactatttcttagaagaggttagcattcaaatcttcaataaatgtg 505
           | | ||     |  | |    |||      || |     ||  |   ||  ||  |
```

```
Db     541 ----GACCTACCGATGCCGTGTGACGGAATGTTCTACCTAGACAACCGCGAAGAACAGGA 596

Qy     506 gacttatagtttggttccaggatggaaaattttgggttaccgttgacctcttca---tgg 562
           |  || |  || |||| || | ||||  || |       | ||||||   | |   |
Db     597 TAAGTACACGTT-GTTCGAGAACGGAACTTTCTTTCGCCACTTTGACCGTGTGACTCTTC 655

Qy     563 aaaaacaggactactgcttgtaccggcacaactttgattcggattttccaaagtccatgt 622
             || | ||| ||||||  |  | |  |  |  || |     | |  | | | | ||| ||
Db     656 GCAAGCGGGAATACTGCCTTCAGCATCTTACATTCGCAGATGGTAATGCTACGTCTATTC 715

Qy     623 ggataatacgacaccgctgcacaagccacatatctcctggatccttagagattctaatta 682
           | ||   ||  |||  |||    |       |||   |           | || | || |
Db     716 GAATTGCACCTCACAACTGTTTGATAGTGCCATCAATTACCGGTCAGACGGTTGTGATGA 775

Qy     683 taa---caatgatatgctttgtcctaacaatcgcagtatatctatacattaagaagctgc 739
           |      |  |||||||| | || ||||| ||||| ||||| || | | | |||||  || |
Db     776 TCTCTTCGCTGATATGCATGGTTCTAACGATCGCCGTATACCTCTTCGTCAAGAAACTAC 835

Qy     740 gaaatgttactggcaagtgcattgtatgctgtatagtttcgaggtttatccagtgcttga 799
           |||  |   ||| || ||| |  | ||||  || || |     |||||    |   |
Db     836 AAAACTTGCATGGAAAATGCTTCATCTGCTACATGGTGTGTCTCTTTATGGGATATCTTT 895

Qy     800 tcatgatactagatcatttaaatctattgaatggcatttgctctccagctggttacagct 859
           || |  | || |||   |   |  |||  | | || |||||   ||||| ||||
Db     896 TCCTATTGCTCGATTTGTGGCAGATATCCATTAGCTTTTGCAAACCAGCAGGTTTTCTGG 955

Qy     860 cgcacttttttccggatggcttccaacctctggctctccgtcatcagctaccatacgtgga 919
           |||| ||   |||||  |      | |||| ||||||||||||    ||    ||||
Db     956 GTTACTTCTTTGTCATGGCCGCATTTTTTTGGCTTTCCGTCATCAGTCTGCACCTTTGGA 1015

Qy     920 aagtcttgacgtcgctc------------aatcgagtcgaccctaactatcggttcctgc 967
           |   || |        |            |||||  ||     | |  |||||||  |||
Db    1016 ACACGTTCAGAGGCTCCTCCCACAAAGCGAATCGCTTCTTATTTGAGCATCGGTTTCTGG 1075

Qy     968 ggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagtatttatatagtta 1027
           |||||  |||  |  ||| |||  || |   || |||| |||| ||    | | |
Db    1076 CCTACAATACCTATGCTTGGGGCATGGCGGTGGTCCTGACAGGAATTACCGTTCTGGCCG 1135

Qy    1028 atcagatttgggaaaacgatcccagtaaatggaactggttgcctctggtcggttttattc 1087
           || | ||   |||||| |      | ||             | | ||| |         |
Db    1136 ATAACATCGTGGAAAACCAGGATTGGAAT---------CCTCGTGTGGGCCACGAGGGAC 1186

Qy    1088 ggtgctcggtcaaagattggcacccatccgtctggatctatataagtggaccgtcgctgg 1147
           |  | | |  |   |    |   || | |  | |||| |    ||  ||   | |
Db    1187 ACTGTTGGATATATACTCAAGCCTGGTCAGCCATGCTCTACTTTTACGGTCCAATGGTAT 1246

Qy    1148 ccctgagcactttcaatgtcgccatgttcgccctgacagccatttacattaggaaagtga 1207
           || |   | || ||  |  ||||||||  |||||| || |     ||      |||||
Db    1247 TTCTTATTGCCTTTAACATAACCATGTTCATCCTGACGGCTAAGCGTATATTAGGAGTGA 1306

Qy    1208 agggggggtataaataagttcacaaatgaggaggagggaaggattaa 1253
           ||  ||  ||  | || ||  |  |  | | ||| ||| |   ||
Db    1307 AGAAGGACATTCAGAACTTTGCGCACAGGCAAGAGAGAAAGCAGAA 1352
```

```
RESULT  14
AC015222
LOCUS       AC015222   143393 bp    DNA             HTG       16-NOV-1999
DEFINITION  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
            pieces.
ACCESSION   AC015222
VERSION     AC015222.1  GI:6436113
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 143393)
  AUTHORS   Adams,M. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
```

```
COMMENT     This sequence was identified as CDM:10211810 by the submitter.
            For further information on this sequence e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .143393
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
BASE COUNT    40751 a  31102 c  30428 g  41112 t
ORIGIN
```

Query Match 6.2%; Score 94.8; DB 63; Length 143393;
Best Local Similarity 49.5%; Pred. No. 3.9e-15;
Matches 373; Conservative 0; Mismatches 307; Indels 74; Gaps 2;

```
Qy     835 atttgctctccagctggttacagctcgcacttttccggatggcttccaacctctggctc 894
           | ||  | | | || ||||| | |  |  ||| ||     | ||         | ||| ||
Db  101585 ACTTTATTTTCCGCAGGTTATATCATGTTCTTCTTTGATTTTGCCAATTCTGTTTGGTTC 101644

Qy     895 tccgtcatcagctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgacgctaac 954
           ||| | ||||||||     |  |||||     | || ||  ||  | |    ||   | |
Db  101645 TCCATAATCAGCTATTGCATATGGAAGAAAATCACATCTGTCGTTAGCCAAGAAAATCGC 101704

Qy     955 tatcggttcctgcggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagt 1014
           ||| |||| |    ||||   | || | ||   | | | | || ||        ||||
Db  101705 GATCAGTTCGTCTTTTACAGTACTTTTGCCTACGGGATATCTGCTATACCACTCGGAATC 101764

Qy    1015 atttatatagttaatcagatttgggaaaacgatcccagtaaatggaactggttgcctctg 1074
           ||   |    |||||||  |||||||| | |||| ||| ||||||||||||||||  |
Db  101765 ATCATTTCGATTAATCAATTTTGGGAAGAGGATCTCAGGAAATGGAACTGGTTGCCCTTA 101824

Qy    1075 gtcggttttattcggtgctcggtcaa---------------------------------- 1100
           || ||||||| | | ||||||||  |
Db  101825 GTTGGTTTTAGTAGATGCTCGGTGGATGGTACGTTTCGATTGGGGAAGACATTTAATGTA 101884

Qy    1101 --------------------------------agattggcacccatccgtctggatctat 1128
                                           ||||||  | |  ||   |||| | |||
Db  101885 TTAATTTAATTCATATATTTACTCCAATTCTCAGATTGTAAGCGGTCTTCCTGGGTTTAT 101944

Qy    1129 ataagtggaccgtcgctggccctgagcactttcaatgtcgccatgttcgccctgacagcc 1188
           |  ||  || |    |  | || |  |  | ||| |    |||||||  |||||   |
Db  101945 TATTTTGTGCCCTATGCGATCATGTGTGCCATTAATATAATTATGTTCGTTCTGACTATC 102004

Qy    1189 atttacattaggaaagtgaaggggggtataaataagttcacaaa--------tgaggagg 1240
           |   |||||| ||||   ||  |   | |  | || || |||||          ||||
Db  102005 AAACACATTATGAAAACAAAAAGAAATCTCCACAATTTAACAAAACGACCAGATAGGAAC 102064

Qy    1241 agggaaggattaactgcataaactttgacagccagacgtaagatatgataaaaccttgga 1300
           |    |  | |||| | |   |||  |            |  | |  |||  | | |
Db  102065 GAGACTTGTGTTACTGTAAATTTTTTAAATTTTGAGTTGTAAGTTTCTTAACGCTTGGAG 102124

Qy    1301 ccttggtaaactagttatccttgccgtattatccacatagttacctacagttcctgcggc 1360
           | ||       | | | |     |   |   |  | ||| ||||||  ||| ||||||
Db  102125 CTTTTTATTCTTTGATTTAAAATTCCATTGCCCTCCTTAGGTACCTATTGTTTCTGCGGA 102184

Qy    1361 tctccatcgtgatgggccttacttggatattcaatgtcattccgtattctgcacggctcc 1420
           | |||    | |||||| |  |||||||| | |   |  ||  ||     | |  |
Db  102185 TTTCCGGTATTATGGGCGTAGCTTGGATACTGATCATTTTTTTGTTGATAGAAGTGAATT 102244

Qy    1421 acattttctgggaatgggtcgggataatatccgagtattttcacagcgcgtttggaattg 1480
           || ||| |||||     |||| ||||||  | |   |  | ||     |  ||||| ||
Db  102245 CCACTTTTTGGGACATTTTCGGTATAATAATCCAACAAATACATTATGGTTTTGGTATCA 102304

Qy    1481 ttttgttcgttctgctcgtcctgaagcgcagcac 1514
           |||| || || |||||  |  | ||||||||||
Db  102305 TTTTATTTGTCCTGCTTATTTTTAAGCGCAGCAC 102338
```

```
RESULT 15
AC009343
```

```
LOCUS       AC009343   182494 bp    DNA             INV       14-MAR-2001
DEFINITION  Drosophila melanogaster, chromosome 2R, region 47A-47B, BAC clone
            BACR08L17, complete sequence.
ACCESSION   AC009343
VERSION     AC009343.7  GI:13324736
KEYWORDS    HTG.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 182494)
  AUTHORS   Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
            Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
            Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
            Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
            Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
            Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
            Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
            Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
            McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
            Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
            Phouanenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
            Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
            Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
  TITLE     Sequencing of Drosophila chromosome 2R, region 47A-47B
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 182494)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
            Rubin,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
            Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT     On Mar 14, 2001 this sequence version replaced gi:6838838.
            Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory, MS 64-121
            Berkeley, CA 94720
            This sequence was assembled using end sequences from a whole genome
            shotgun and from subclones of this BAC and its neighboring clones.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive Web site (http://www.fruitfly.org/sequence/) or send email
            to bdgp@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .182494
                     /organism="Drosophila melanogaster"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /chromosome="2R"
                     /map="47A-47B"
                     /clone="BACR08L17 (D1050)"
                     /clone_lib="RPCI-98 (Roswell Park Cancer Institute
                     Drosophila melanogaster BAC library, partial EcoRI in
                     pBACe3.6)"
BASE COUNT    53541 a  38776 c  38094 g  52083 t
ORIGIN
```

Query Match 6.2%; Score 94.8; DB 4; Length 182494;
Best Local Similarity 49.5%; Pred. No. 4e-15;
Matches 373; Conservative 0; Mismatches 307; Indels 74; Gaps 2;

```
Qy     835 atttgctctccagctggttacagctcgcacttttccggatggcttccaacctctggctc 894
           | ||  | | | || ||||| | |  |  ||| ||     | ||         | ||| ||
Db  159176 ACTTTATTTTCCGCAGGTTATATCATGTTCTTCTTTGATTTTGCCAATTCTGTTTGGTTC 159235
```

```
Qy     895 tccgtcatcagctaccatacgtggaaagtcttgacgtcgctcaatcgagtcgaccctaac 954
           ||| | ||||||||     |  |||||     | || ||  ||  | |    ||   | |
Db 159236 TCCATAATCAGCTATTGCATATGGAAGAAAATCACATCTGTCGTTAGCCAAGAAAATCGC 159295

Qy     955 tatcggttcctgcggtacaacgccttcgtctggagcacagccgcaatcatgacgggaagt 1014
           ||| |||| |    ||||   | || | ||   | | | | || ||       ||||
Db 159296 GATCAGTTCGTCTTTTACAGTACTTTTGCCTACGGGATATCTGCTATACCACTCGGAATC 159355

Qy    1015 atttatatagttaatcagatttgggaaaacgatcccagtaaatggaactggttgcctctg 1074
           ||   |    |||||||  |||||||| | |||| ||| ||||||||||||||||  |
Db 159356 ATCATTTCGATTAATCAATTTTGGGAAGAGGATCTCAGGAAATGGAACTGGTTGCCCTTA 159415

Qy    1075 gtcggttttattcggtgctcggtcaa---------------------------------- 1100
           || ||||||| | | ||||||||  |
Db 159416 GTTGGTTTTAGTAGATGCTCGGTGGATGGTACGTTTCGATTGGGGAAGACATTTAATGTA 159475

Qy    1101 --------------------------------agattggcacccatccgtctggatctat 1128
                                            ||||||  | |  ||   |||| | |||
Db 159476 TTAATTTAATTCATATATTTACTCCAATTCTCAGATTGTAAGCGGTCTTCCTGGGTTTAT 159535

Qy    1129 ataagtggaccgtcgctggccctgagcactttcaatgtcgccatgttcgccctgacagcc 1188
             ||  || |    |  | || |  |  | ||| |    |||||||  |||||   |
Db 159536 TATTTTGTGCCCTATGCGATCATGTGTGCCATTAATATAATTATGTTCGTTCTGACTATC 159595

Qy    1189 atttacattaggaaagtgaaggggggtataaataagttcacaaa--------tgaggagg 1240
           |   |||||| ||||   ||  |   | |  | || || |||||          ||||
Db 159596 AAACACATTATGAAAACAAAAAGAAATCTCCACAATTTAACAAAACGACCAGATAGGAAC 159655

Qy    1241 agggaaggattaactgcataaactttgacagccagacgtaagatatgataaaaccttgga 1300
             |    |  | |||| | |   ||| |            |  | |  |||  | | |
Db 159656 GAGACTTGTGTTACTGTAAATTTTTTAAATTTTGAGTTGTAAGTTTCTTAACGCTTGGAG 159715

Qy    1301 ccttggtaaactagttatccttgccgtattatccacatagttacctacagttcctgcggc 1360
           | ||       | | | |      |   |   |  | ||| ||||||  ||| ||||||
Db 159716 CTTTTTATTCTTTGATTTAAAATTCCATTGCCCTCCTTAGGTACCTATTGTTTCTGCGGA 159775

Qy    1361 tctccatcgtgatgggccttacttggatattcaatgtcattccgtattctgcacggctcc 1420
           | |||    | |||||| |  |||||||| | |   |  ||  ||     | |  |
Db 159776 TTTCCGGTATTATGGGCGTAGCTTGGATACTGATCATTTTTTTGTTGATAGAAGTGAATT 159835

Qy    1421 acattttctgggaatgggtcgggataatatccgagtattttcacagcgcgtttggaattg 1480
           || ||| |||||     |||| ||||||  | |  |  | ||    |  ||||| ||
Db 159836 CCACTTTTTGGGACATTTTCGGTATAATAATCCAACAAATACATTATGGTTTTGGTATCA 159895

Qy    1481 ttttgttcgttctgctcgtcctgaagcgcagcac 1514
           |||| || || |||||  |  | ||||||||||
Db 159896 TTTTATTTGTCCTGCTTATTTTTAAGCGCAGCAC 159929


Search completed: October 30, 2001, 11:43:40
Job time: 4422 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2001, 10:52:33 ; Search time 122.38 Seconds
(without alignments)
7880.837 Million cell updates/sec

Title: US-09-619-049-854
Perfect score: 1536
Sequence: 1 atgaggcttccttgggtgat.....ggactctaatgatggattct 1536

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601:*
1: /SIDS1/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	90	5.9	936	22	AAF58252	Oligonucleotide D1
c	2	90	5.9	936	22	AAF58254	Oligonucleotide D1
c	3	90	5.9	936	22	AAF58257	Oligonucleotide D1
c	4	90	5.9	936	22	AAF58259	Oligonucleotide D2
c	5	90	5.9	936	22	AAF58262	Oligonucleotide D2
c	6	90	5.9	938	22	AAF58255	Oligonucleotide D1
	7	89.8	5.8	936	22	AAF58252	Oligonucleotide D1
	8	89.8	5.8	936	22	AAF58254	Oligonucleotide D1
	9	89.8	5.8	936	22	AAF58257	Oligonucleotide D1
	10	89.8	5.8	936	22	AAF58259	Oligonucleotide D2
	11	89.8	5.8	936	22	AAF58262	Oligonucleotide D2
	12	89.8	5.8	938	22	AAF58255	Oligonucleotide D1
	13	42.6	2.8	244	22	AAF58238	Oligonucleotide D1
c	14	40.4	2.6	244	22	AAF58238	Oligonucleotide D1
c	15	38.4	2.5	6516	19	AAV52238	Streptococcus pneu
c	16	35.4	2.3	1664976	19	AAV21209	Methanococcus jann
c	17	34.8	2.3	777	20	AAZ16530	Human gene express
	18	34.8	2.3	1573	22	AAF72752	Human prostate can
c	19	34.8	2.3	141589	21	AAF20913	Human ELAM-1 polyn
c	20	34.8	2.3	141589	21	AAF21127	Human low adenosin
c	21	34.8	2.3	141589	21	AAF21152	Human low adenosin
c	22	34.8	2.3	141589	21	AAA35005	Human adenosine re
c	23	34.8	2.3	141589	21	AAA35030	Human adenosine re
c	24	34.8	2.3	146981	21	AAF21442	Human ELAM-1 polyn
c	25	34.8	2.3	209273	21	AAF21437	Human factor-relat
	26	34.2	2.2	273254	21	AAC81914	Chlamydia pneumoni
	27	34	2.2	1401	18	AAT73167	Pseudomonas putida
	28	34	2.2	1412	18	AAT73175	Pseudomonas putida
	29	34	2.2	1441	18	AAT73170	Pseudomonas putida
c	30	34	2.2	1685	21	AAA26702	Candida albicans p
c	31	34	2.2	1998	20	AAZ20986	Human semaphorin Z
	32	33.6	2.2	1031	20	AAX97702	Extended human sec
c	33	33.6	2.2	1111	19	AAX14395	H. pylori GHPO 561
	34	33.4	2.2	1417	18	AAT73171	Pseudomonas putida
	35	33.4	2.2	6216	19	AAV06946	Canine herpes viru
c	36	33.4	2.2	26281	21	AAZ60929	Nucleotide sequenc
c	37	33.2	2.2	1233	21	AAC51237	Arabidopsis thalia
c	38	33.2	2.2	1336	21	AAC38025	Arabidopsis thalia
c	39	33.2	2.2	1341	21	AAC51247	Arabidopsis thalia
	40	33.2	2.2	1664976	19	AAV21209	Methanococcus jann
c	41	33	2.1	138	16	AAT26436	Human gene signatu
c	42	33	2.1	151826	21	AAF22291	BAC containing rep
c	43	33	2.1	269223	22	AAF28554	Genomic fragment #
c	44	32.6	2.1	1750	19	AAV23688	L.lactis DB1341 or
c	45	32.4	2.1	667	20	AAX79036	Human secreted pro

ALIGNMENTS

RESULT 1
AAF58252/c
ID AAF58252 standard; DNA; 936 BP.
XX
AC AAF58252;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1835.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX

```
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;


  Query Match             5.9%;  Score 90;  DB 22;  Length 936;
  Best Local Similarity   0.8%;  Pred. No. 3.7e-18;
  Matches    6;  Conservative  460;  Mismatches  320;  Indels    0;  Gaps    0;

Qy      46 tttacgaacaactcaaatgcagatattcccgggtgcaactactacgacacggttgatatc 105
           :::: :: :: : :::: : :::::     :  :: :: ::  : :   :: ::::
Db     787 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 728

Qy     106 tcatacattgaaaggcaaaacgattcgtatttatatgatgacatcgaaattcctgctagc 165
           : ::: ::: :::   :::: :::  :::::::::  : : ::  ::::: :  ::
Db     727 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWW 668

Qy     166 ttaactggatactacgagttcaggcagtttggggacggttcgattacgccgattgaaaag 225
           :::: :  ::: ::  : :: :   : :::     :  ::  ::::     ::: ::::
Db     667 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 608

Qy     226 catttaagggcctgtgtttgcagcgtgcggccctgtattcgaatctgctgcccagccaag 285
           ::::::     : : :::  :   :        : ::::  ::: :  :    :   ::
Db     607 WWWWWWWWWWWWWWWCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 548

Qy     286 aacttttggccaacggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaa 345
           :: :::::    ::   :::::   :: ::  : : ::: :  :  :      |:: :::
Db     547 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWW 488

Qy     346 ccctatatatacttcacatacatggacctacaggcacgagtaccactcaccgatatggct 405
              :::::::: :: : ::: ::  :  :: :   :  : ::  : : :   ::::   :
Db     487 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 428

Qy     406 attatcagagatgagttctttgattgtgatgaaatgatttacatcagcgactttaactat 465
           ::::: : : :: : :: ::: ::: : ::  ||  ::::: :: :   : ::::: :::
Db     427 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWTAAGCWWWWWWWWWWWWWWWWWWWWWWW 368

Qy     466 ttcttagaagaggttagcattcaaatcttcaataaatgtggacttatagtttggttccag 525
           :: ::: :: :  :::  ::: :::: :: ::::::: :  : ::::: :::  ::  :
Db     367 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 308

Qy     526 gatggaaaattttgggttaccgttgacctcttcatggaaaaacaggactactgcttgtac 585
            ::  :::::::::|  :::   :: :  : :: ::  ::::: :  : :: :  :: ::
Db     307 WWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 248

Qy     586 cggcacaactttgattcggattttccaaagtccatgtggataatacgacaccgctgcaca 645
           : :: ::: :::   :::::  ::: :  :: : :::::  : : |  :  : :
Db     247 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWCWWWWWW 188

Qy     646 agccacatatctcctggatccttagagattctaattataacaatgatatgctttgtccta 705
           :   : :::: :  :  ::  ::: : ::: ::::::::: ::: ::::  ::: :  ::
Db     187 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 128

Qy     706 acaatcgcagtatatctatacattaagaagctgcgaaatgttactggcaagtgcattgta 765
           : :::   : ::::: :::: ::::: ::| :   :::: ::: :   :: :  ::: ::
Db     127 WWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWW 68

Qy     766 tgctgtatagtttcgaggtttatccagtgcttgatcatgatactagatcatttaaatcta 825
           :  : :::: ::: :  :::::  : :  :: :: :: ::: :: :: :::::::: ::
Db      67 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 8

Qy     826 ttgaat 831
           :: :::
Db       7 WWWWWW 2
```

```
RESULT   2
AAF58254/c
ID   AAF58254 standard; DNA; 936 BP.
XX
AC   AAF58254;
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D1875.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;


  Query Match             5.9%;  Score 90;  DB 22;  Length 936;
  Best Local Similarity   0.8%;  Pred. No. 3.7e-18;
  Matches    6;  Conservative  460;  Mismatches  320;  Indels    0;  Gaps    0;

Qy      46 tttacgaacaactcaaatgcagatattcccgggtgcaactactacgacacggttgatatc 105
           :::: :: :: : :::: : :::::     :  :: :: ::  : :   :: ::::
Db     787 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 728

Qy     106 tcatacattgaaaggcaaaacgattcgtatttatatgatgacatcgaaattcctgctagc 165
           : ::: ::: :::   :::: :::  :::::::::  : : ::  ::::: :  ::
Db     727 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWW 668

Qy     166 ttaactggatactacgagttcaggcagtttggggacggttcgattacgccgattgaaaag 225
           :::: :  ::: ::  : :: :   : :::     :  ::  ::::     ::: ::::
Db     667 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 608

Qy     226 catttaagggcctgtgtttgcagcgtgcggccctgtattcgaatctgctgcccagccaag 285
           ::::::     : : :::  :   :        : ::::  ::: :  :    :   ::
Db     607 WWWWWWWWWWWWWWWCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 548

Qy     286 aacttttggccaacggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaa 345
           :: :::::    ::   :::::   :: ::  : : ::: :  :  :      |:: :::
Db     547 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWW 488

Qy     346 ccctatatatacttcacatacatggacctacaggcacgagtaccactcaccgatatggct 405
              :::::::: :: : ::: ::  :  :: :   :  : ::  : : :   ::::   :
Db     487 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 428
```

```
Qy    406 attatcagagatgagttctttgattgtgatgaaatgatttacatcagcgactttaactat 465
          ::::: : : :: : :: ::: ::: : ::  ||  ::::: :: :   : ::::: :::
Db    427 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWTAAGCWWWWWWWWWWWWWWWWWWWWWWW 368

Qy    466 ttcttagaagaggttagcattcaaatcttcaataaatgtggacttatagtttggttccag 525
          :: ::: :: :  :::  ::: :::: :: :::::: :  : ::::: :::  ::  :
Db    367 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 308

Qy    526 gatggaaaattttgggttaccgttgacctcttcatggaaaaacaggactactgcttgtac 585
           ::  ::::::::|  :::   :: :  : :: ::  ::::: :  : :: :  :: ::
Db    307 WWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 248

Qy    586 cggcacaactttgattcggattttccaaagtccatgtggataatacgacaccgctgcaca 645
          : :: ::: :::   :::::  ::: :  :: :  ::::::  : : |  :  : :
Db    247 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWCWWWWWWWW 188

Qy    646 agccacatatctcctggatccttagagattctaattataacaatgatatgctttgtccta 705
          :   : :::: :  :  ::  ::: : ::: ::::::::: ::: ::::  ::: :  ::
Db    187 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 128

Qy    706 acaatcgcagtatatctatacattaagaagctgcgaaatgttactggcaagtgcattgta 765
          : :::   : ::::: :::: ::::: ::| :   :::: ::: :   :: :  ::: ::
Db    127 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWW 68

Qy    766 tgctgtatagtttcgaggtttatccagtgcttgatcatgatactagatcatttaaatcta 825
          :  : :::: :::  :  :::::  : :  :: :: :: ::: :: :: :::::::: ::
Db     67 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 8

Qy    826 ttgaat 831
          :: :::
Db      7 WWWWWW 2


RESULT   3
AAF58257/c
ID   AAF58257 standard; DNA; 936 BP.
XX
AC   AAF58257;
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D1954.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface
XX
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;


  Query Match             5.9%;  Score 90;  DB 22;  Length 936;
  Best Local Similarity   0.8%;  Pred. No. 3.7e-18;
  Matches    6;  Conservative  460;  Mismatches  320;  Indels    0;  Gaps    0;

Qy     46 tttacgaacaactcaaatgcagatattcccgggtgcaactactacgacacggttgatatc 105
          ::::  :: :: : ::::  : :::::      :  :: :: ::  : :  :: ::::
Db    787 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 728

Qy    106 tcatacattgaaaggcaaaacgattcgtatttatatgatgacatcgaaattcctgctagc 165
          : ::: ::: :::   ::::  :::  :::::::::  : : ::  :::::  :  ::
Db    727 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWW 668

Qy    166 ttaactggatactacgagttcaggcagtttggggacggttcgattacgccgattgaaaag 225
          :::: :  ::: ::  : :: :   : :::    :  ::  ::::     ::: ::::
Db    667 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 608

Qy    226 catttaagggcctgtgtttgcagcgtgcggccctgtattcgaatctgctgcccagccaag 285
           ::::::    : : :::  :  :  :  : ::::  ::: : :  :  :  ::
Db    607 WWWWWWWWWWWWWWWWCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 548

Qy    286 aacttttttggccaacggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaa 345
          :: :::::    ::   :::::  :: ::  : : ::: :  :  :      |:: :::
Db    547 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWW 488

Qy    346 ccctatatatacttcacatacatggacctacaggcacgagtaccactcaccgatatggct 405
             :::::::: :: : ::: ::  :  :: :   :  : ::  : : :   ::::   :
Db    487 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 428

Qy    406 attatcagagatgagttctttgattgtgatgaaatgatttacatcagcgactttaactat 465
          ::::: : : :: : :: ::: ::: : ::  ||  ::::: :: :   : ::::: :::
Db    427 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWTAAGCWWWWWWWWWWWWWWWWWWWWWWW 368

Qy    466 ttcttagaagaggttagcattcaaatcttcaataaatgtggacttatagtttggttccag 525
          :: ::: :: :  :::  ::: :::: :: :::::: :  : ::::: :::  ::  :
Db    367 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 308

Qy    526 gatggaaaattttgggttaccgttgacctcttcatggaaaaacaggactactgcttgtac 585
           ::  ::::::::|  :::   :: :  : :: ::  ::::: :  : :: :  :: ::
Db    307 WWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 248

Qy    586 cggcacaactttgattcggattttccaaagtccatgtggataatacgacaccgctgcaca 645
          : :: ::: :::   :::::  ::: :  :: :  ::::::  : : |  :  : :
Db    247 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWCWWWWWWWW 188

Qy    646 agccacatatctcctggatccttagagattctaattataacaatgatatgctttgtccta 705
          :   : :::: :  :  ::  ::: : ::: ::::::::: ::: ::::  ::: :  ::
Db    187 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 128

Qy    706 acaatcgcagtatatctatacattaagaagctgcgaaatgttactggcaagtgcattgta 765
          : :::   : ::::: :::: ::::: ::| :   :::: ::: :   :: :  ::: ::
Db    127 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWW 68

Qy    766 tgctgtatagtttcgaggtttatccagtgcttgatcatgatactagatcatttaaatcta 825
          :  : :::: :::  :  :::::  : :  :: :: :: ::: :: :: :::::::: ::
Db     67 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 8

Qy    826 ttgaat 831
          :: :::
Db      7 WWWWWW 2


RESULT   4
AAF58259/c
ID   AAF58259 standard; DNA; 936 BP.
XX
AC   AAF58259;
```

```
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D2004.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 128; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;

  Query Match             5.9%;  Score 90;  DB 22;  Length 936;
  Best Local Similarity   0.8%;  Pred. No. 3.7e-18;
  Matches    6;  Conservative  460;  Mismatches  320;  Indels    0;  Gaps    0;

Qy      46 tttacgaacaactcaaatgcagatattcccgggtgcaactactacgacacggttgatatc 105
           ::::  :: :: : ::::  : :::::      :  :: :: ::  : :  :: ::::
Db     787 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 728

Qy     106 tcatacattgaaaggcaaaacgattcgtatttatatgatgacatcgaaattcctgctagc 165
           : ::: ::: :::   ::::  :::  ::::::::  : : ::  ::::: :  ::
Db     727 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWW 668

Qy     166 ttaactggatactacgagttcaggcagtttggggacggttcgattacgccgattgaaaag 225
           :::: :  ::: ::  : :: :   : :::     :   ::  ::::     ::: ::::
Db     667 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 608

Qy     226 catttaagggcctgtgtttgcagcgtgcggccctgtattcgaatctgctgcccagccaag 285
            ::::::     : : :::  :   :        : ::::  ::: : :   :   ::
Db     607 WWWWWWWWWWWWWWWCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 548

Qy     286 aacttttggccaacggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaa 345
           :: :::::    ::   :::::   :: ::  : : ::: :  :  :      |:: :::
Db     547 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWW 488

Qy     346 ccctatatatacttcacatacatggacctacaggcacgagtaccactcaccgatatggct 405
              :::::::: :: : ::: ::   :  :: :   :  : ::  : : :   ::::   :
Db     487 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 428

Qy     406 attatcagagatgagttctttgattgtgatgaaatgatttacatcagcgactttaactat 465
           ::::: : : :: : :: ::: ::: : ::  ||  ::::: :: :   : ::::: :::
Db     427 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWTAAGCWWWWWWWWWWWWWWWWWWWW 368
```

```
Qy     466 ttcttagaagaggttagcattcaaatcttcaataaatgtggacttatagtttggttccag 525
           :: ::: :: :  :::  ::: :::: :: ::::::: :  : ::::: :::  ::  :
Db     367 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 308

Qy     526 gatggaaaattttgggttaccgttgacctcttcatggaaaaacaggactactgcttgtac 585
            ::  ::::::::|  :::   :: :  : :: ::  ::::: :  : :: :  :: ::
Db     307 WWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 248

Qy     586 cggcacaactttgattcggattttccaaagtccatgtggataatacgacaccgctgcaca 645
              : :: ::: :::   :::::  ::: :  :: :  ::::::  : : |  :  : :
Db     247 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWCWWWWW 188

Qy     646 agccacatatctcctggatccttagagattctaattataacaatgatatgctttgtccta 705
           :   : :::: :  :  ::  ::: : ::: :::::::: ::: :::  ::: :  ::
Db     187 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 128

Qy     706 acaatcgcagtatatctatacattaagaagctgcgaaatgttactggcaagtgcattgta 765
           : :::   : ::::: :::: ::::: ::| :   :::: ::: :   :: :  ::: ::
Db     127 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWW 68

Qy     766 tgctgtatagtttcgaggtttatccagtgcttgatcatgatactagatcatttaaatcta 825
           :  : :::: :::  :  ::::  : :  :: :: :: ::: :: :: :::::::: ::
Db      67 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 8

Qy     826 ttgaat 831
           :: :::
Db       7 WWWWWW 2
```

```
RESULT   5
AAF58262/c
ID   AAF58262 standard; DNA; 936 BP.
XX
AC   AAF58262;
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D2007.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 128; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 5 A; 139 C; 10 G; 6 T; 776 other;
```

```
  Query Match             5.9%;  Score 90;  DB 22;  Length 936;
  Best Local Similarity   0.8%;  Pred. No. 3.7e-18;
  Matches    6;  Conservative  460;  Mismatches  320;  Indels    0;  Gaps    0;

Qy      46 tttacgaacaactcaaatgcagatattcccgggtgcaactactacgacacggttgatatc 105
           ::::  :: :: : ::::  : :::::      :  :: :: ::  : :   :: ::::
Db     787 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 728

Qy     106 tcatacattgaaaggcaaaacgattcgtatttatatgatgacatcgaaattcctgctagc 165
           : ::: ::: :::   :::: :::  :::::::::  : : ::  ::::: :  ::
Db     727 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWW 668

Qy     166 ttaactggatactacgagttcaggcagtttggggacggttcgattacgccgattgaaaag 225
           :::: :  ::: ::  : :: :   : :::    :   ::  ::::     ::: ::::
Db     667 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 608

Qy     226 catttaagggcctgtgtttgcagcgtgcggccctgtattcgaatctgctgcccagccaag 285
           ::::::     : : :::  :   :        : ::::  ::: :  :    :   ::
Db     607 WWWWWWWWWWWWWWWWCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 548

Qy     286 aacttttggccaacggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaa 345
           :: :::::    ::   :::::   :: ::  : : ::: :  :  :       |:: :::
Db     547 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWW 488

Qy     346 ccctatatatacttcacatacatggacctacaggcacgagtaccactcaccgatatggct 405
           :::::::: :: : ::: ::  :  :: :  :  : ::  :  : :   ::::  :
Db     487 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 428

Qy     406 attatcagagatgagttctttgattgtgatgaaatgatttacatcagcgactttaactat 465
           ::::: : : :: : :: ::: ::: : ::  ||  ::::: :: :   : ::::: :::
Db     427 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWTAAGCWWWWWWWWWWWWWWWWWWW 368

Qy     466 ttcttagaagaggttagcattcaaatcttcaataaatgtggacttatagtttggttccag 525
           :: ::: :: :  :::  ::: :::: :: ::::::: :  : ::::: :::  ::  :
Db     367 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 308

Qy     526 gatggaaaattttgggttaccgttgacctcttcatggaaaaacaggactactgcttgtac 585
            ::  ::::::::|  :::   :: :  : :: ::  ::::: :  : :: :  :: ::
Db     307 WWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 248

Qy     586 cggcacaactttgattcggattttccaaagtccatgtggataatacgacaccgctgcaca 645
           : :: ::: :::   ::::  ::: :  :: :  ::::::  : : |  :  : :
Db     247 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWCWWWWWWWW 188

Qy     646 agccacatatctcctggatccttagagattctaattataacaatgatatgctttgtccta 705
           :   : :::: :  :  ::  ::: : ::: :::::::: ::: ::::  ::: :  ::
Db     187 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 128

Qy     706 acaatcgcagtatatctatacattaagaagctgcgaaatgttactggcaagtgcattgta 765
           : :::   : ::::: :::  :::::  : :  :: :: :: :: :: :: :::::: ::
Db     127 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWW 68

Qy     766 tgctgtatagtttcgaggtttatccagtgcttgatcatgatactagatcatttaaatcta 825
           :  : :::: :::  :  :::::  : :  :: :: :: :: :: :: :: :::::::: ::
Db      67 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 8

Qy     826 ttgaat 831
           :: :::
Db       7 WWWWWW 2


RESULT   6
AAF58255/c
ID   AAF58255 standard; DNA; 938 BP.
XX
AC   AAF58255;
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D1876.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface
XX
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;


  Query Match             5.9%;  Score 90;  DB 22;  Length 938;
  Best Local Similarity   0.8%;  Pred. No. 3.7e-18;
  Matches    6;  Conservative  460;  Mismatches  320;  Indels    0;  Gaps    0;

Qy      46 tttacgaacaactcaaatgcagatattcccgggtgcaactactacgacacggttgatatc 105
           ::::  :: :: : ::::  : :::::      :  :: :: ::  : :   :: ::::
Db     787 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 728

Qy     106 tcatacattgaaaggcaaaacgattcgtatttatatgatgacatcgaaattcctgctagc 165
           : ::: ::: :::   :::: :::  :::::::::  : : ::  ::::: :  ::
Db     727 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWW 668

Qy     166 ttaactggatactacgagttcaggcagtttggggacggttcgattacgccgattgaaaag 225
           :::: :  ::: ::  : :: :   : :::    :   ::  ::::     ::: ::::
Db     667 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 608

Qy     226 catttaagggcctgtgtttgcagcgtgcggccctgtattcgaatctgctgcccagccaag 285
           ::::::     : : :::  :   :        : ::::  ::: :  :    :   ::
Db     607 WWWWWWWWWWWWWWWWCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 548

Qy     286 aacttttggccaacggaaaatgcgatgatggtctcaaagaggagctcgcccggttcaaa 345
           :: :::::    ::   :::::   :: ::  : : ::: :  :  :       |:: :::
Db     547 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWW 488

Qy     346 ccctatatatacttcacatacatggacctacaggcacgagtaccactcaccgatatggct 405
           :::::::: :: : ::: ::  :  :: :  :  : ::  :  : :   ::::  :
Db     487 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 428

Qy     406 attatcagagatgagttctttgattgtgatgaaatgatttacatcagcgactttaactat 465
           ::::: : : :: : :: ::: ::: : ::  ||  ::::: :: :   : ::::: :::
Db     427 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWTAAGCWWWWWWWWWWWWWWWWWWW 368

Qy     466 ttcttagaagaggttagcattcaaatcttcaataaatgtggacttatagtttggttccag 525
           :: ::: :: :  :::  ::: :::: :: ::::::: :  : ::::: :::  ::  :
Db     367 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 308

Qy     526 gatggaaaattttgggttaccgttgacctcttcatggaaaaacaggactactgcttgtac 585
```

```
Db    307 WWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 248

Qy    586 cggcacaactttgattcggattttccaaagtccatgtggataatacgacaccgctgcaca 645

Db    247 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWCWWWWWWWW 188

Qy    646 agccacatatctcctggatccttagagattctaattataacaatgatatgctttgtccta 705

Db    187 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 128

Qy    706 acaatcgcagtatatctatacattaagaagctgcgaaatgttactggcaagtgcattgta 765

Db    127 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 68

Qy    766 tgctgtatagtttcgaggtttatccagtgcttgatcatgatactagatcatttaaaatcta 825

Db     67 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 8

Qy    826 ttgaat 831

Db      7 WWWWWW 2
```

RESULT 7
AAF58252
ID AAF58252 standard; DNA; 936 BP.
XX
AC AAF58252;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1835.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX
PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match 5.8%; Score 89.8; DB 22; Length 936;
Best Local Similarity 0.9%; Pred. No. 4.2e-18;
Matches 7; Conservative 441; Mismatches 303; Indels 0; Gaps 0;

```
Qy     98 ttgatatctcatacattgaaaggcaaaacgattcgtatttatatgatgacatcgaaattc 157

Db     39 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwc 98

Qy    158 ctgctagcttaactggatactacgagttcaggcagtttggggacggttcgattacgccga 217

Db     99 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 158

Qy    218 ttgaaaagcatttaagggcctgtgtttgcagcgtgcggccctgtattcgaatctgctgcc 277

Db    159 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgwwwwwwwwwwwwwwwwwwww 218

Qy    278 cagccaagaacttttggccaacggaaaatgcgatgatggtctcaaagaggagctcgccc 337

Db    219 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 278

Qy    338 ggttcaaaccctatatatacttcacatacatggacctacaggcacgagtaccactcaccg 397

Db    279 wwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 338

Qy    398 atatggctattatcagagatgagttctttgattgtgatgaaatgatttacatcagcgact 457

Db    339 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgcttaww 398

Qy    458 ttaactatttcttagaagaggttagcattcaaatcttcaataaatgtggacttatagttt 517

Db    399 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 458

Qy    518 ggttccaggatggaaaattttgggttaccgttgacctcttcatggaaaaacaggactact 577

Db    459 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwww 518

Qy    578 gcttgtaccggcacaactttgattcggattttccaaagtccatgtggataatacgacacc 637

Db    519 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 578

Qy    638 gctgcacaagccacatatctcctggatccttagagattctaattataacaatgatatgct 697

Db    579 wwwwwwwwwwwwwgwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 638

Qy    698 ttgtcctaacaatcgcagtatatctatacattaagaagctgcgaaatgttactggcaagt 757

Db    639 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwww 698

Qy    758 gcattgtatgctgtatagtttcgaggtttatccagtgcttgatcatgatactagatcatt 817

Db    699 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 758

Qy    818 taaatctattgaatggcatttgctctccagc 848

Db    759 wwwwwwwwwwwwwwwwwwwwwwwwwwwwgc 789
```

RESULT 8
AAF58254
ID AAF58254 standard; DNA; 936 BP.
XX
AC AAF58254;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1875.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX

```
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

  Query Match             5.8%;  Score 89.8;  DB 22;  Length 936;
  Best Local Similarity   0.9%;  Pred. No. 4.2e-18;
  Matches    7;  Conservative  441;  Mismatches  303;  Indels    0;  Gaps    0;

Qy      98 ttgatatctcatacattgaaaggcaaaacgattcgtatttatatgatgacatcgaaattc 157
           :: :::: : ::: ::: :::    ::::  :::  ::::::::: :: : ::  :::::|
Db      39 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwc 98

Qy     158 ctgctagcttaactggatactacgagttcaggcagtttggggacggttcgattacgccga 217
             :  ::  :::: :  ::: ::  : :: :    : :::     :   ::  ::::     :
Db      99 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 158

Qy     218 ttgaaaagcatttaagggcctgtgtttgcagcgtgcggccctgtattcgaatctgctgcc 277
           :: ::::  ::::::     : : :::  :   :   |  : ::::  ::: :  :
Db     159 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgwwwwwwwwwwwwwwwwwwwwww 218

Qy     278 cagccaagaacttttggccaacggaaaatgcgatgatggtctcaaagaggagctcgccc 337
             :   :: :: :::::    ::    :::::   :: ::  : : ::: :  :  :
Db     219 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 278

Qy     338 ggttcaaaccctatatatacttcacatacatggacctacaggcacgagtaccactcaccg 397
              :: :::   :::: ::: :: : ::: ::   :  :: :   :  : ::  : : :
Db     279 wwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 338

Qy     398 atatggctattatcagagatgagttctttgattgtgatgaaatgatttacatcagcgact 457
           ::::   :::::: : : :: : :: ::: ::: : :: :::: ::::: ::      | :
Db     339 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgcttaww 398

Qy     458 ttaactatttcttagaagaggttagcattcaaatcttcaataaatgtggacttatagttt 517
           :::: ::::: ::: :: :  :::  ::: :::: :: ::::::: :  : ::::: :::
Db     399 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 458

Qy     518 ggttccaggatggaaaatttgggttaccgttgacctcttcatggaaaaacaggactact 577
              ::  :  ::  ::::::::   :::    :: : |: :: ::  ::::: :  : :: :
Db     459 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwww 518

Qy     578 gcttgtaccggcacaactttgattcggattttccaaagtccatgtggataatacgacacc 637
              :: ::      : :: ::: :::   :::::  ::: :   :: :  ::::::  : :
Db     519 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 578

Qy     638 gctgcacaagccacatatctcctggatccttagagattctaattataacaatgatatgct 697
              :  : ::   : :::: :  :  ::  ::: : ::: ::::::::: ::: ::::  :
Db     579 wwwwwwwwwwwwwgwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 638

Qy     698 ttgtcctaacaatcgcagtatatctatacattaagaagctgcgaaatgttactggcaagt 757
           :: :  ::: :::   : ::::: :::: ::::: ::  :   :::: :::|:   :: :
Db     639 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwww 698

Qy     758 gcattgtatgctgtatagtttcgaggtttatccagtgcttgatcatgatactagatcatt 817
              ::: :::  : :::: :::  :  :::::  : :  :: :: :: ::: :: :: :::
Db     699 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 758

Qy     818 taaatctattgaatggcatttgctctccagc 848
            ::::: :::: :::   ::::  : :  :||
Db     759 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwgc 789

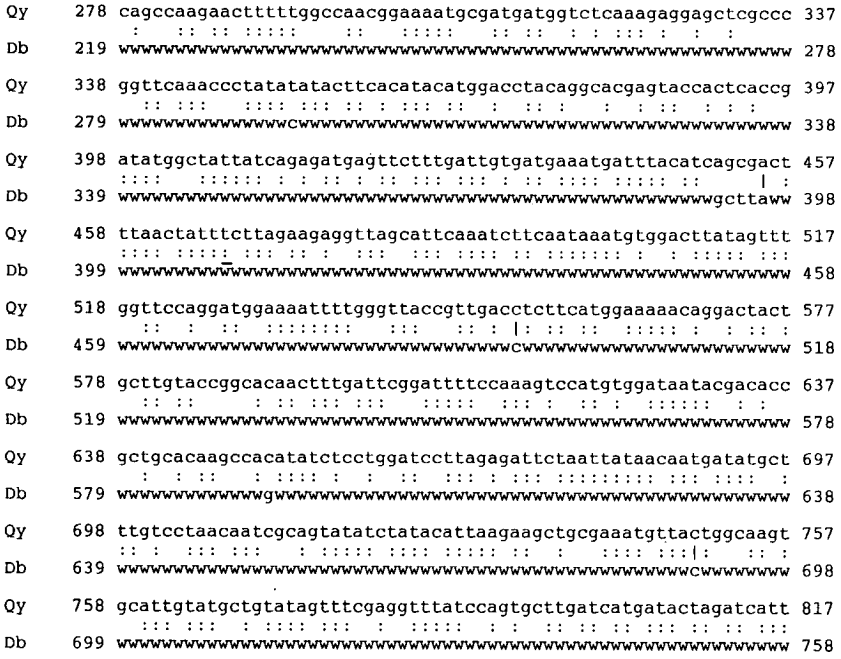

RESULT   9
AAF58257
ID   AAF58257 standard; DNA; 936 BP.
XX
AC   AAF58257;
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D1954.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

  Query Match             5.8%;  Score 89.8;  DB 22;  Length 936;
  Best Local Similarity   0.9%;  Pred. No. 4.2e-18;
  Matches    7;  Conservative  441;  Mismatches  303;  Indels    0;  Gaps    0;

Qy      98 ttgatatctcatacattgaaaggcaaaacgattcgtatttatatgatgacatcgaaattc 157
           :: :::: : ::: ::: :::    ::::  :::  ::::::::: :: : ::  :::::|
Db      39 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwc 98

Qy     158 ctgctagcttaactggatactacgagttcaggcagtttggggacggttcgattacgccga 217
             :  ::  :::: :  ::: ::  : :: :    : :::     :   ::  ::::     :
Db      99 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 158

Qy     218 ttgaaaagcatttaagggcctgtgtttgcagcgtgcggccctgtattcgaatctgctgcc 277
           :: ::::  ::::::     : : :::  :   :   |  : ::::  ::: :  :
Db     159 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgwwwwwwwwwwwwwwwwwwwwww 218
```



RESULT 10
AAF58259
ID AAF58259 standard; DNA; 936 BP.
XX
AC AAF58259;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D2004.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
PS Example 6; Page 128; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;

Query Match 5.8%; Score 89.8; DB 22; Length 936;
Best Local Similarity 0.9%; Pred. No. 4.2e-18;
Matches 7; Conservative 441; Mismatches 303; Indels 0; Gaps 0;

RESULT 11
AAF58262

```
ID   AAF58262 standard; DNA; 936 BP.
XX
AC   AAF58262;
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D2007.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 128; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 936 BP; 5 A; 139 C; 10 G; 6 T; 776 other;

  Query Match             5.8%;  Score 89.8;  DB 22;  Length 936;
  Best Local Similarity   0.9%;  Pred. No. 4.2e-18;
  Matches    7;  Conservative  441;  Mismatches  303;  Indels    0;  Gaps    0;

Qy      98 ttgatatctcatacattgaaaggcaaaacgattcgtatttatatgatgacatcgaaattc 157
              :: :::: : ::: ::: :::    ::::  :::  :::::::: :: : ::  :::::|
Db      39 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwc 98

Qy     158 ctgctagcttaactggatactacgagttcaggcagtttggggacggttcgattacgccga 217
              :  ::  :::: :  ::: ::  : :: :   : :::    :   ::  ::::     :
Db      99 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 158

Qy     218 ttgaaaagcatttaagggcctgtgtttgcagcgtgcggccctgtattcgaatctgctgcc 277
              :: ::::  ::::::     : : :::  :   :   |   : ::::  ::: :  :
Db     159 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgwwwwwwwwwwwwwwwwwwwww 218

Qy     278 cagccaagaactttttggccaacggaaaatgcgatgatggtctcaaagaggagctcgccc 337
              :  :: :: :::::     ::    :::::   :: ::  : : ::: :  :  :
Db     219 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 278

Qy     338 ggttcaaaccctatatatacttcacatacatggacctacaggcacgagtaccactcaccg 397
                :: :::    :::: ::: :: : ::: ::   :: :    :  : ::  : : :
Db     279 wwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 338

Qy     398 atatggctattatcagagatgagttctttgattgtgatgaaatgatttacatcagcgact 457
              ::::   :::::: : : :: : :: ::: ::: : :: :::: ::::: ::     | :
Db     339 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwgcttaww 398

Qy     458 ttaactatttcttagaagaggttagcattcaaatcttcaataaatgtggacttatagttt 517
              :::: ::::: ::: :: :   :::   ::: :::: :: :::::::  :   : ::::: :::
Db     399 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 458

Qy     518 ggttccaggatggaaaatttgggttaccgttgacctcttcatggaaaaacaggactact 577
                :: :  ::  ::::::::   :::    :: : |: :: ::  ::::: :  : :: :
Db     459 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwwwwwwwwwwwwwwwwww 518

Qy     578 gcttgtaccggcacaactttgattcggattttccaaagtccatgtggataatacgacacc 637
                :: ::      : :: ::: :::    :::::  ::: :   :: :  :::::: : :
Db     519 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 578

Qy     638 gctgcacaagccacatatctcctggatccttagagattctaattataacaatgatatgct 697
                : . : ::    : :::: :   :   ::   ::: : ::: ::::::::: ::: ::::   :
Db     579 wwwwwwwwwwwwgwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 638

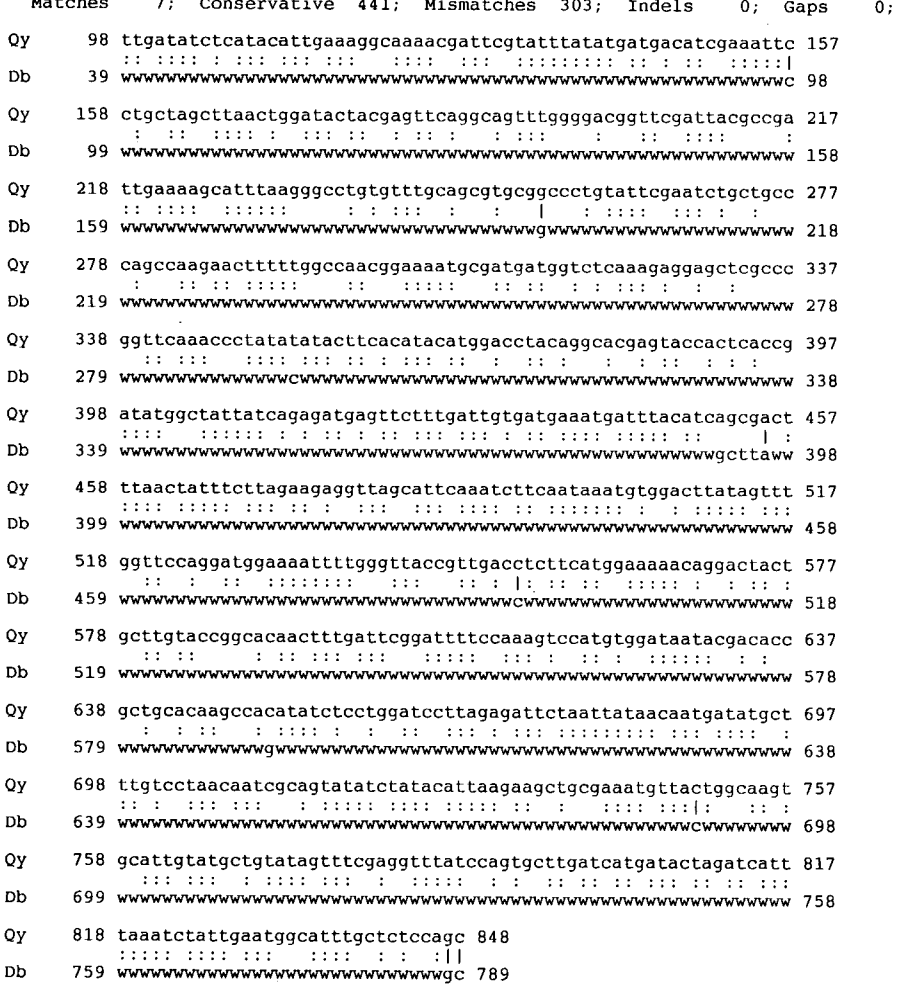
Qy     698 ttgtcctaacaatcgcagtatatctatacattaagaagctgcgaaatgttactggcaagt 757
              :: :   ::: :::    : ::::: :::: ::::: ::   :     :::: :::|:    :: :
Db     639 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwcwwwwwww 698

Qy     758 gcattgtatgctgtatagtttcgaggtttatccagtgcttgatcatgatactagatcatt 817
                ::: :::   : :::: :::   :   :::::    : :   :: :: :: ::: :: :: :::
Db     699 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 758

Qy     818 taaatctattgaatggcatttgctctccagc 848
              ::::: :::: :::    ::::  : :  :||
Db     759 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwgc 789


RESULT 12
AAF58255
ID   AAF58255 standard; DNA; 938 BP.
XX
AC   AAF58255;
XX
DT   24-APR-2001  (first entry)
XX
DE   Oligonucleotide D1876.
XX
KW   Electron-transfer group; ETM; mismatch; genotyping;
KW   gene expression; ss.
XX
OS   Synthetic.
XX
PN   WO200107665-A2.
XX
PD   01-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20476.
XX
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface  -
XX
PS   Example 6; Page 127; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;
```

Query Match 5.8%; Score 89.8; DB 22; Length 938;
Best Local Similarity 0.9%; Pred. No. 4.2e-18;
Matches 7; Conservative 441; Mismatches 303; Indels 0; Gaps 0;



RESULT 13
AAF58238
ID AAF58238 standard; DNA; 244 BP.
XX
AC AAF58238;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1250:D1102.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
PS Example 4; Page 120; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 244 BP; 19 A; 9 C; 12 G; 10 T; 194 other;

Query Match 2.8%; Score 42.6; DB 22; Length 244;
Best Local Similarity 9.0%; Pred. No. 0.0018;
Matches 21; Conservative 124; Mismatches 88; Indels 0; Gaps 0;

RESULT 14
AAF58238/c
ID AAF58238 standard; DNA; 244 BP.
XX
AC AAF58238;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1250:D1102.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX

```
PR   26-JUL-1999;   99US-0145695.
PR   17-MAR-2000; 2000US-0190259.
XX
PA   (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI   Umek RM;
XX
DR   WPI; 2001-159728/16.
XX
PT   Nucleic acids containing electron-transfer group, useful as labels in
PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT   a single surface -
XX
PS   Example 4; Page 120; 159pp; English.
XX
CC   The present invention relates to a composition comprising two nucleic
CC   acids each containing an electron-transfer group (ETM) having
CC   different redox potentials. The invention is used for electronic
CC   detection of nucleic acids, especially of substitutions (mismatches)
CC   and single-nucleotide polymorphisms, e.g. for genotyping,
CC   monitoring gene expression.
XX
SQ   Sequence 244 BP; 19 A; 9 C; 12 G; 10 T; 194 other;


  Query Match             2.6%;  Score 40.4;  DB 22;  Length 244;
  Best Local Similarity   8.5%;  Pred. No. 0.009;
  Matches   17;  Conservative  111;  Mismatches   72;  Indels    0;  Gaps    0;

Qy     639 ctgcacaagccacatatctcctggatccttagagattctaattataacaatgatatgctt 698
           |||||    || |||   |  ||  |  | | | ::: ::::::::: ::: ::::  ::
Db     235 CTGCAGCTTCCTCATTGATGGTGTCTTTTAACAWWWWWWWWWWWWWWWWWWWWWWWWWWW 176

Qy     699 tgtcctaacaatcgcagtatatctatacattaagaagctgcgaaatgttactggcaagtg 758
           : :  ::: :::   : ::::: :::: ::::: ::  :   :::: ::: :   :: :
Db     175 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 116

Qy     759 cattgtatgctgtatagtttcgaggtttatccagtgcttgatcatgatactagatcattt- 818
           ::: :::    :::: :::  :  :::::  : :  :: :: :: ::: :: :: ::::
Db     115 WWWWWWWWWWGWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 56

Qy     819 aaatctattgaatggcattt 838
           :::: :::: :::   ::::
Db      55 WWWWWWWWWWWWWWWWWWWW 36


RESULT  15
AAV52238/c
ID   AAV52238 standard; DNA; 6516 BP.
XX
AC   AAV52238;
XX
DT   23-OCT-1998  (first entry)
XX
DE   Streptococcus pneumoniae genome fragment SEQ ID NO:105.
XX
KW   Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KW   computer readable medium; vaccine; pharmaceutical composition; ds.
XX
OS   Streptococcus pneumoniae.
XX
PN   WO9818931-A2.
XX
PD   07-MAY-1998.
XX
PF   30-OCT-1997;   97WO-US19588.
XX
PR   31-OCT-1996;   96US-0029960.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Choi GH,  Dillon PJ,  Dougherty BA,  Fannon M;
PI   Kunsch CA,  Rosen CA;
XX
DR   WPI; 1998-272225/24.
XX
PT   Computer-readable medium with recorded Streptococcus pneumoniae
PT   polynucleotide sequences - useful in diagnostic kits and assays, and
PT   pharmaceutical compositions and vaccines for Streptococcus
PT   pneumoniae
XX
PS   Claim 1; Page 777-781; 1409pp; English.
XX
CC   The present invention describes a computer readable medium which has
CC   the nucleotide sequences SEQ ID NO:1 to 391. (AAV52134 to AAV52524)
CC   recorded on it, or a representative fragment or a sequence at least 95%
CC   identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in
CC   SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
CC   Streptococcus pneumoniae. The present invention also describes an
CC   isolated nucleic acid molecule encoding a homologue of any of the
CC   fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the
CC   nucleic acid molecule is produced by a process comprising: (a) screening
CC   a genomic DNA library using as a probe a target sequence defined by any
CC   of the sequences in SEQ ID NO:1 to 391, identifying members of the
CC   library which contain sequences that hybridise to the target sequence and
CC   isolating the nucleic acid molecules from the members; or (b) isolating
CC   mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
CC   molecules whose nucleotide sequence is homologous to amplification
CC   primers derived from the fragment of the S. pneumoniae genome to prime
CC   the amplification and isolating the amplified sequences. The computer
CC   readable medium can be used in a computer-based system for identifying
CC   fragments of the S. pneumoniae genome of commercial importance, or
CC   expression modulating fragments of the S. pneumoniae genome. Products
CC   from the present invention can be used in diagnosis kits and assays, and
CC   pharmaceutical compositions and vaccines for S. pneumoniae.
XX
SQ   Sequence 6516 BP; 1882 A; 1438 C; 1259 G; 1935 T; 2 other;


  Query Match             2.5%;  Score 38.4;  DB 19;  Length 6516;
  Best Local Similarity   54.2%;  Pred. No. 0.29;
  Matches   78;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy     336 ccggttcaaaccctatatatacttcacatacatggacctacaggcacgagtaccactcac 395
           | ||  ||||    | |  | | | | |  | |  ||||  | ||| ||   |  |  |||
Db    3796 CTGGGACAAAGAGGAGAGGTTCCTGATCTGCTAGGACAAAGAGGAACATTTGGCGTTCAT 3737

Qy     396 cgatatggctattatcagagatgagttctttgattgtgatgaaatgatttacatcagcga 455
           |  ||  || | |  |    |||| |  |||||||||| |   | |  |||  ||||||
Db    3736 GGGGATCACTGTAAAAACCAGTGAGGTAATTGATTGTGACGGGGTCAGATACGACAGCGA 3677

Qy     456 ctttaactatttcttagaagaggt 479
           | |  | | ||||| |    || |
Db    3676 CGTCTAGTTTTTCTGATTCAAGAT 3653


Search completed: October 30, 2001, 12:07:47
Job time: 4514 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2001, 10:28:18 ; Search time 53.9 Seconds
(without alignments)
5394.840 Million cell updates/sec

Title: US-09-619-049-854
Perfect score: 1536
Sequence: 1 atgaggcttccttgggtgat.....ggactctaatgatggattct 1536

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	34	2.2	1401	4	US-08-952-457-1	Sequence 1, Appli
	2	33.4	2.2	6216	3	US-09-213-053-1	Sequence 1, Appli
	3	32.4	2.1	7218	1	US-08-232-463-14	Sequence 14, Appl
c	4	32	2.1	511	1	US-08-412-614-97	Sequence 97, Appl
c	5	32	2.1	511	1	US-08-412-614-98	Sequence 98, Appl
c	6	32	2.1	511	2	US-08-635-761-97	Sequence 97, Appl
c	7	32	2.1	511	2	US-08-635-761-98	Sequence 98, Appl
	8	31.4	2.0	1554	1	US-08-469-486-1	Sequence 1, Appli
	9	31.4	2.0	1554	2	US-08-469-658-1	Sequence 1, Appli
	10	31.2	2.0	534	3	US-08-714-918-28	Sequence 28, Appl
	11	31.2	2.0	534	4	US-09-265-315-28	Sequence 28, Appl
	12	31.2	2.0	534	4	US-09-265-315-28	Sequence 28, Appl
	13	31.2	2.0	534	4	US-09-266-417-28	Sequence 28, Appl
	14	31.2	2.0	1365	1	US-08-114-072-1	Sequence 1, Appli
	15	31.2	2.0	1365	5	PCT-US94-09361-1	Sequence 1, Appli
c	16	31.2	2.0	9636	1	US-08-323-170B-1	Sequence 1, Appli
c	17	31	2.0	2400	1	US-08-487-753-15	Sequence 15, Appl
c	18	31	2.0	2400	2	US-08-480-065-15	Sequence 15, Appl
c	19	31	2.0	2400	3	US-08-487-744-15	Sequence 15, Appl
	20	30.8	2.0	969	2	US-08-932-978-1	Sequence 1, Appli
c	21	30.6	2.0	9606	1	US-07-741-940-1	Sequence 1, Appli
c	22	30.6	2.0	9606	1	US-08-289-548A-1	Sequence 1, Appli
c	23	30.6	2.0	9606	1	US-08-452-654-1	Sequence 1, Appli
c	24	30.6	2.0	9606	2	US-08-370-235A-1	Sequence 1, Appli
	25	30.4	2.0	51952	3	US-08-947-823-1	Sequence 1, Appli
	26	30.4	2.0	87350	3	US-08-781-891-79	Sequence 79, Appl
	27	30.2	2.0	1952	2	US-08-392-806A-5	Sequence 5, Appli
	28	30.2	2.0	1952	4	US-09-257-490-5	Sequence 5, Appli
c	29	30	2.0	696	2	US-08-336-198C-6	Sequence 6, Appli
	30	30	2.0	1125	1	US-07-928-462-1	Sequence 1, Appli
	31	30	2.0	1125	3	US-08-273-247-1	Sequence 1, Appli
	32	30	2.0	2483	4	US-08-894-324A-3	Sequence 3, Appli
c	33	30	2.0	3959	1	US-08-474-067-1	Sequence 1, Appli
c	34	30	2.0	3959	2	US-08-474-068A-1	Sequence 1, Appli
c	35	30	2.0	3959	2	US-08-472-481-1	Sequence 1, Appli
c	36	30	2.0	4771	2	US-08-866-650-2	Sequence 2, Appli
c	37	30	2.0	4771	2	US-09-021-287-2	Sequence 2, Appli
	38	30	2.0	43676	3	US-09-356-952-12	Sequence 12, Appl
c	39	29.8	1.9	5156	2	US-09-091-432-3	Sequence 3, Appli
c	40	29.6	1.9	7101	1	US-08-480-604A-9	Sequence 9, Appli
c	41	29.6	1.9	7101	2	US-08-405-496A-9	Sequence 9, Appli
	42	29.4	1.9	1279	3	US-08-961-083-47	Sequence 47, Appl
c	43	29.4	1.9	7824	5	PCT-US95-06722-11	Sequence 11, Appl
	44	29.2	1.9	1926	1	US-07-718-535-5	Sequence 5, Appli
	45	29.2	1.9	1926	1	US-08-161-999-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-952-457-1
; Sequence 1, Application US/08952457A
; Patent No. 6251655
; GENERAL INFORMATION:
; APPLICANT: Minambres RODRIGUEZ, Baltasar
; APPLICANT: Martinez BLANCO, Honorina
; APPLICANT: Robriguez OLIVERA, Elias
; APPLICANT: Garcia ALONSO, Belen
; APPLICANT: Fernandez CANON, Jose Manuel
; APPLICANT: Barredo FUENTE, Jose Luis
; APPLICANT: Diez GARCIA, Bruno
; APPLICANT: Schleissner SANCHEZ, Carmen
; APPLICANT: Moreno VALLE, Miguel Angel
; APPLICANT: Salto MALDONADO, Francisco
; APPLICANT: Luengo RODRIGUEZ, Jose Maria
; TITLE OF INVENTION: PROCESS FOR INCREASING THE PRODUCTION OF PENICILLIN G
; TITLE OF INVENTION: (BENZYLPENICILLIN) IN PENICILLUM CHRYSOGENUM BY
; TITLE OF INVENTION: EXPRESSION OF THE PCL GENE
; FILE REFERENCE: U-011501-7
; CURRENT APPLICATION NUMBER: US/08/952,457A
; CURRENT FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: PCT/ES97/00069
; EARLIER FILING DATE: 1997-03-18
; EARLIER APPLICATION NUMBER: P9600664
; EARLIER FILING DATE: 1996-03-18
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1401
; TYPE: DNA
; ORGANISM: Pseudomonas
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: The sequence
; OTHER INFORMATION: is a fusion of a DNA sequence originating from
; OTHER INFORMATION: pUC18 with another originating from Pseudomonas
; OTHER INFORMATION: corresponding to the pcl gene.rdp
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)..(1386)
US-08-952-457-1

Query Match 2.2%; Score 34; DB 4; Length 1401;
Best Local Similarity 54.9%; Pred. No. 0.35;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 863 acttttttccggatggcttccaacctctggctctccgtcatcagctaccatacgtggaaag 922
Db 5 acactcaccgcgtggcttgcaaccgctggcgcgcggcgtacaagaacaattcgagtgaag 64

```
Qy    923 tcttgacgtcgctcaatcgagtcgaccctaactatcggttcctgcggtacaacgccttcg 982
          | |||    |  | ||   | |||||  |  |     |  || |  | | | ||||   |
Db     65 ccatgaacatgtaccatgatgccgaccgtgccctgttggacccgatggaaaccgccagtg 124

Qy    983 tc 984
          ||
Db    125 tc 126


RESULT   2
US-09-213-053-1
; Sequence 1, Application US/09213053
; Patent No. 6159477
; GENERAL INFORMATION:
;  APPLICANT: AUDONNET, Jean-Christophe
;  APPLICANT: BAUDU, Philippe
;  TITLE OF INVENTION: CANINE HERPESVIRUS BASED RECOMBINANT LIVE VACCINE, IN
;  TITLE OF INVENTION: PARTICULAR AGAINST CANINE DISTEMPER, RABIES OR THE
;  TITLE OF INVENTION: PARAINFLUENZA 2 VIRUS
;  FILE REFERENCE: 454313-2200
;  CURRENT APPLICATION NUMBER: US/09/213,053
;  CURRENT FILING DATE: 1998-12-16
;  EARLIER APPLICATION NUMBER: 9608242
;  EARLIER FILING DATE: 1996-06-27
;  EARLIER APPLICATION NUMBER: PCT/FR97/01115
;  EARLIER FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 6216
;   TYPE: DNA
;   ORGANISM: Canine herpesvirus
US-09-213-053-1


  Query Match             2.2%;  Score 33.4;  DB 3;  Length 6216;
  Best Local Similarity   52.5%;  Pred. No. 1.6;
  Matches   73;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy    595 tttgattcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacata 654
          || |  | | | ||   |||   ||  |  |||| |        ||  ||||   ||| |
Db   2438 ttagtcttgaacttaataaatatcaattaaataaaatactcgaactaaacaaaaaacaaa 2497

Qy    655 tctcctggatccttagagattctaattataacaatgatatgctttgtcctaacaatcgca 714
            || | | |       || |||||| | || |||| | || |||   || | | | | |
Db   2498 attcattgttaaaatttgaatctaatgaaaataatgttgtgatttcatctgaaagtggaa 2557

Qy    715 gtatatctatacattaaga 733
          || | ||  |  |||  ||
Db   2558 gtgtttcattgaatttgga 2576


RESULT   3
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14


  Query Match             2.1%;  Score 32.4;  DB 1;  Length 7218;
  Best Local Similarity   4.5%;  Pred. No. 4;
  Matches   18;  Conservative  202;  Mismatches  178;  Indels    0;  Gaps    0;

Qy    513 agtttggttccaggatggaaaattttgggttaccgttgacctcttcatggaaaaacagga 572
          || |||| | ||||  |    |  ||| | | :: ::  ::::::: :       :
Db   1036 AGCTTGGCTGCAGGTCGAGGGAGCTTGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYY 1095

Qy    573 ctactgcttgtaccggcacaactttgattcggatttttccaaagtccatgtggataatacg 632
          :: :: ::: : ::  : :  :::: :::   ::::::    ::: : :   :  : :
Db   1096 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1155

Qy    633 acaccgctgcacaagccacatatctcctggatccttagagattctaattataacaatgat 692
            : :: :: : :   :: : : ::::::  :::::    ::::  :: : :  : : :
Db   1156 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1215

Qy    693 atgctttgtcctaacaatcgcagtatatctatacattaagaagctgcgaaatgttactgg 752
            : :::: ::::  :  :: :  : : ::: : : ::      :: :    : :: ::
Db   1216 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1275

Qy    753 caagtgcattgtatgctgtatagtttcgaggtttatccagtgcttgatcatgatactaga 812
          :   : : :: : : :: : :  ::::    ::: :::  : :::  :: :  : ::
Db   1276 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1335

Qy    813 tcatttaaatctattgaatggcatttgctctccagctggttacagctcgcacttttttccg 872
          :: :::   ::: ::   :  : ::: ::::::  ::  :: :  ::: : ::::::::
Db   1336 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1395

Qy    873 gatggcttccaacctctggctctccgtcatcagctacc 910
            :  :::::  :::::  :::::: :: ::  :: ::
Db   1396 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1433


RESULT   4
US-08-412-614-97/c
; Sequence 97, Application US/08412614
; Patent No. 5536638
;  GENERAL INFORMATION:
;    APPLICANT:  Rossau, Rudi
;    APPLICANT:  Van Heuverswyn, Hugo
;    TITLE OF INVENTION:  Hybridization Probes Derived from the
;    TITLE OF INVENTION:  Spacer Region Between the 16S and 23S rRNA Genes for the
;    TITLE OF INVENTION:  Detection of No. 5536638-Viral Microorganisms
```

```
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant & Gould
;      STREET:  3100 No. 5536638west Center
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  USA
;      ZIP:  55402-4131
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette 3.5 inch, 1.44 mb capacity
;      COMPUTER:  IBM PC compatible (Compaq Deskpro 286e)
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Wordperfect Version #5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/412,614
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/965,394
;      FILING DATE:  17-DEC-1992
;      APPLICATION NUMBER:  PCT/EP91/00743
;      FILING DATE:  18-APR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB/90901054.3
;      FILING DATE:  18-APR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hillson, Randall A.
;      REGISTRATION NUMBER:  31,838
;      REFERENCE/DOCKET NUMBER:  8076.75-USWO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612-332-5300
;      TELEFAX:  612-332-9081
;  INFORMATION FOR SEQ ID NO:  97:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  511 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Campylobacter jejuni
;      STRAIN:  ATCC 33560
US-08-412-614-97


  Query Match             2.1%;  Score 32;  DB 1;  Length 511;
  Best Local Similarity   47.9%;  Pred. No. 0.8;
  Matches   92;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy     662 gatccttagagattctaattataacaatgatatgctttgtcctaacaatcgcagtatatc 721
           ||  ||||   ||  |||||| ||  |  || |   ||        ||  | || ||||
Db     214 GACTCTTAACAATGATAATTAAAAAGAACATTTAGGTTTAAAACCTAAAGGAAGCATATA 155

Qy     722 tatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcga 781
           |||  | |||     |   |||| | | |    | || ||    |  |      | |  |
Db     154 TATTAAATAATCTCATTATAAATATAAATACATAATGTATATGTTATTTATATTTATTAA 95

Qy     782 ggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgct 841
           || ||| |  |  | | | | ||||  |    || ||| | ||  |  ||   |||
Db      94 AGTATATTCTTTTATAGTTAAAGATAAAAATAAATCTAAGTATAGAGTTTGTATTTTATA 35

Qy     842 ctccagctggtt 853
           |  | ||  ||
Db      34 TTTAATCTACTT 23


RESULT   5
US-08-412-614-98/c
; Sequence 98, Application US/08412614
; Patent No. 5536638
;  GENERAL INFORMATION:
;    APPLICANT:  Rossau, Rudi
;    APPLICANT:  Van Heuverswyn, Hugo
;    TITLE OF INVENTION:  Hybridization Probes Derived from the
;    TITLE OF INVENTION:  Spacer Region Between the 16S and 23S rRNA Genes for the
;    TITLE OF INVENTION:  Detection of No. 5536638-Viral Microorganisms
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant & Gould
;      STREET:  3100 No. 5536638west Center
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  USA
;      ZIP:  55402-4131
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette 3.5 inch, 1.44 mb capacity
;      COMPUTER:  IBM PC compatible (Compaq Deskpro 286e)
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Wordperfect Version #5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/412,614
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/965,394
;      FILING DATE:  17-DEC-1992
;      APPLICATION NUMBER:  PCT/EP91/00743
;      FILING DATE:  18-APR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB/90901054.3
;      FILING DATE:  18-APR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hillson, Randall A.
;      REGISTRATION NUMBER:  31,838
;      REFERENCE/DOCKET NUMBER:  8076.75-USWO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612-332-5300
;      TELEFAX:  612-332-9081
;  INFORMATION FOR SEQ ID NO:  98:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  511 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Campylobacter coli
;      STRAIN:  ATCC 33559
US-08-412-614-98


  Query Match             2.1%;  Score 32;  DB 1;  Length 511;
  Best Local Similarity   47.9%;  Pred. No. 0.8;
  Matches   92;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy     662 gatccttagagattctaattataacaatgatatgctttgtcctaacaatcgcagtatatc 721
           ||  ||||   ||  |||||| ||  |  || |   ||        ||  | || ||||
Db     214 GACTCTTAACAATGATAATTAAAAAGAACATTTAGGTTTAAAACCTAAAGGAAGCATATA 155

Qy     722 tatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcga 781
           |||  | |||     |   |||| | | |    | || ||    |  |      | |  |
Db     154 TATTAAATAATCTCATTATAAATATAAATACATAATGTATATGTTATTTATATTTATTAA 95

Qy     782 ggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgct 841
           || ||| |  |  | | | | ||||  |    || ||| | ||  |  ||   |||
Db      94 AGTATATTCTTTTATAGTTAAAGATAAAAATAAATCTAAGTATAGAGTTTGTATTTTATA 35

Qy     842 ctccagctggtt 853
           |  | ||  ||
Db      34 TTTAATCTACTT 23
```

```
RESULT   6
US-08-635-761-97/c
; Sequence 97, Application US/08635761
; Patent No. 5945282
;  GENERAL INFORMATION:
;    APPLICANT:  Rossau, Rudi
;    TITLE OF INVENTION:  HYBRIDIZATION PROBES DERIVED FROM THE SPACER REGION BE
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant, Gould, Smith, Edell, Welter & Schmidt
;      STREET:  3100 No. 5945282west Center, 90 S. 7th Street
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  U.S.A.
;      ZIP:  55402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/635,761
;      FILING DATE:  19-APR-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/965,394
;      FILING DATE:  17-DEC-1992
;      APPLICATION NUMBER:  08/412,614
;      FILING DATE:  29-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hillson, Randall A
;      REGISTRATION NUMBER:  31,838
;      REFERENCE/DOCKET NUMBER:  8076.75USC1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612/332-5300
;      TELEFAX:  612/332/9081
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  97:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  511 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
US-08-635-761-97


  Query Match             2.1%;  Score 32;  DB 2;  Length 511;
  Best Local Similarity   47.9%;  Pred. No. 0.8;
  Matches   92;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy     662 gatccttagagattctaattataacaatgatatgctttgtcctaacaatcgcagtatatc 721
           ||  ||||   ||  |||||| ||  |  || |   ||        ||  | || ||||
Db     214 GACTCTTAACAATGATAATTAAAAAGAACATTTAGGTTTAAAACCTAAAGGAAGCATATA 155

Qy     722 tatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcga 781
           |||  | |||     |   |||| | | |    | || ||    |  |       | |  |
Db     154 TATTAAATAATCTCATTATAAATATAAATACATAATGTATATGTTATTTATATTTATTAA 95

Qy     782 ggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgct 841
           || ||| |  |  | | | | ||||  |    || ||| | ||  |  ||   |||
Db      94 AGTATATTCTTTTATAGTTAAAGATAAAAATAAATCTAAGTATAGAGTTTGTATTTTATA 35

Qy     842 ctccagctggtt 853
           |  | ||  ||
Db      34 TTTAATCTACTT 23


RESULT   7
```

```
US-08-635-761-98/c
; Sequence 98, Application US/08635761
; Patent No. 5945282
;  GENERAL INFORMATION:
;    APPLICANT:  Rossau, Rudi
;    TITLE OF INVENTION:  HYBRIDIZATION PROBES DERIVED FROM THE SPACER REGION BE
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant, Gould, Smith, Edell, Welter & Schmidt
;      STREET:  3100 No. 5945282west Center, 90 S. 7th Street
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  U.S.A.
;      ZIP:  55402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/635,761
;      FILING DATE:  19-APR-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/965,394
;      FILING DATE:  17-DEC-1992
;      APPLICATION NUMBER:  08/412,614
;      FILING DATE:  29-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hillson, Randall A
;      REGISTRATION NUMBER:  31,838
;      REFERENCE/DOCKET NUMBER:  8076.75USC1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612/332-5300
;      TELEFAX:  612/332/9081
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  98:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  511 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
US-08-635-761-98


  Query Match             2.1%;  Score 32;  DB 2;  Length 511;
  Best Local Similarity   47.9%;  Pred. No. 0.8;
  Matches   92;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy     662 gatccttagagattctaattataacaatgatatgctttgtcctaacaatcgcagtatatc 721
           ||  ||||   ||  |||||| ||  |  || |   ||        ||  | || ||||
Db     214 GACTCTTAACAATGATAATTAAAAAGAACATTTAGGTTTAAAACCTAAAGGAAGCATATA 155

Qy     722 tatacattaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcga 781
           |||  | |||     |   |||| | | |    | || ||    |  |       | |  |
Db     154 TATTAAATAATCTCATTATAAATATAAATACATAATGTATATGTTATTTATATTTATTAA 95

Qy     782 ggtttatccagtgcttgatcatgatactagatcatttaaatctattgaatggcatttgct 841
           || ||| |  |  | | | | ||||  |    || ||| | ||  |  ||   |||
Db      94 AGTATATTCTTTTATAGTTAAAGATAAAAATAAATCTAAGTATAGAGTTTGTATTTTATA 35

Qy     842 ctccagctggtt 853
           |  | ||  ||
Db      34 TTTAATCTACTT 23


RESULT   8
US-08-469-486-1
```

```
; Sequence 1, Application US/08469486
; Patent No. 5739281
;  GENERAL INFORMATION:
;    APPLICANT:  Thoegersen, Hans Christian
;    APPLICANT:  Holtet, Thor Las
;    APPLICANT:  Etzerodt, Michael
;    TITLE OF INVENTION:  Improved method for the refolding of
;    TITLE OF INVENTION:  proteins
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version
;      SOFTWARE:  #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,486
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/192,060
;      FILING DATE:  February 4, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  06363/002001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617 542 5070
;      TELEFAX:  617 542 8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1554 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Bos taurus
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  76..1551
US-08-469-486-1


  Query Match             2.0%;  Score 31.4;  DB 1;  Length 1554;
  Best Local Similarity   49.1%;  Pred. No. 2.9;
  Matches   83;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy   1115 ccgtctggatctatataagtggaccgtcgctggccctgagcactttcaatgtcgccatgt 1174
           || ||  || | |  |    | |   ||   ||||   ||     | | | |   | ||
Db    863 CCATCCTGAACGAGTTCTACGTCCTCACGGCTGCCCACTGCCTGCACCAGGCCAAGAGGT 922

Qy   1175 tcgccctgacagccatttacattaggaaagtgaaggggggtataaataagttcacaaatg 1234
           || |  |||  | |    ||   |  | || | ||| ||     ||  || |  || | |
Db    923 TCACGGTGAGGGTCGGCGACCGGAACACAGAGCAGGAGGAGGGCAACGAGATGGCACACG 982

Qy   1235 aggaggagggaaggattaactgcataaactttgacagccagacgtaaga 1283
           ||| ||||   |  | ||  ||    ||||| ||   |||| || ||
Db    983 AGGTGGAGATGACTGTGAAGCACAGCCGCTTTGTCAAGGAGACCTACGA 1031


RESULT   9
US-08-469-658-1
```

```
; Sequence 1, Application US/08469658
; Patent No. 5917018
;  GENERAL INFORMATION:
;    APPLICANT:  Th egersen, Hans Christian
;    APPLICANT:  Holtet, Thor Las
;    APPLICANT:  Etzerodt, Michael
;    TITLE OF INVENTION:  IMPROVED METHOD FOR THE REFOLDING OF
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version
;      SOFTWARE:  #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,658
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/192,060
;      FILING DATE:  February 4, 1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  06363/002002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617 542 5070
;      TELEFAX:  617 542 8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1554 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Bos taurus
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  76..1551
US-08-469-658-1


  Query Match             2.0%;  Score 31.4;  DB 2;  Length 1554;
  Best Local Similarity   49.1%;  Pred. No. 2.9;
  Matches   83;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy   1115 ccgtctggatctatataagtggaccgtcgctggccctgagcactttcaatgtcgccatgt 1174
           || ||  || | |  |    | |   ||   ||||   ||     | | | |   | ||
Db    863 CCATCCTGAACGAGTTCTACGTCCTCACGGCTGCCCACTGCCTGCACCAGGCCAAGAGGT 922

Qy   1175 tcgccctgacagccatttacattaggaaagtgaaggggggtataaataagttcacaaatg 1234
           || |  |||  | |    ||   |  | || | ||| ||     ||  || |  || | |
Db    923 TCACGGTGAGGGTCGGCGACCGGAACACAGAGCAGGAGGAGGGCAACGAGATGGCACACG 982

Qy   1235 aggaggagggaaggattaactgcataaactttgacagccagacgtaaga 1283
           ||| ||||   |  | ||  ||    ||||| ||   |||| || ||
Db    983 AGGTGGAGATGACTGTGAAGCACAGCCGCTTTGTCAAGGAGACCTACGA 1031


RESULT  10
```

```
US-08-714-918-28
; Sequence 28, Application US/08714918
; Patent No. 6037123
;  GENERAL INFORMATION:
;    APPLICANT:  Benton, Bret
;    APPLICANT:  Lee, Ving
;    APPLICANT:  Malouin, Francois
;    APPLICANT:  Martin, Patrick K.
;    APPLICANT:  Schmid, Molly B.
;    APPLICANT:  Sun, Dongxu
;    TITLE OF INVENTION:  STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
;    TITLE OF INVENTION:  TARGET GENES
;    NUMBER OF SEQUENCES:  111
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/714,918
;      FILING DATE:  September 13, 1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/009,102
;      FILING DATE:  December 22, 1995
;      APPLICATION NUMBER:  60/003,798
;      FILING DATE:  September 15, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  222/005
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  534 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-714-918-28


  Query Match             2.0%;  Score 31.2;  DB 3;  Length 534;
  Best Local Similarity   50.3%;  Pred. No. 1.5;
  Matches   72;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy   1387 atattcaatgtcattccgtattctgcacggctccacattttctgggaatgggtcgggata 1446
          |||   || ||  | || |  | |   ||||||  |  | | |       ||| | | |
Db     30 ATAACTAAGGTGGTGCCATTATGTNAGCGGCTCGTCCATGTNTATTTGGCGGTAGTGGTT 89

Qy   1447 atatccgagtattttcacagcgcgtttggaattgttttgttcgttctgctcgtcctgaag 1506
            |||  |||  | |||| |         |||      | | ||||  | | || ||||
Db     90 TAATCACTGTAGCTACACATCCAACAAATCATTCAGCATTACATTCTCGTNGCCCAGAAG 149

Qy   1507 cgcagcacatggactctaatgat 1529
          ||  |   || ||   |||||||
Db    150 CGATGGTTATTGATATTAATGAT 172


RESULT  11
US-09-265-315-28
; Sequence 28, Application US/09265315
```

```
; Patent No. 6187541
;  GENERAL INFORMATION:
;    APPLICANT:  Benton, Bret
;    APPLICANT:  Lee, Ving J.
;    APPLICANT:  Malouin, Francois
;    APPLICANT:  Martin, Patrick K.
;    APPLICANT:  Schmid, Molly B.
;    APPLICANT:  Sun, Dongxu
;    TITLE OF INVENTION:  METHODS OF SCREENING FOR COMPOUNDS
;    TITLE OF INVENTION:  ACTIVE ON STAPHYLOCOCCUS AUREUS
;    TITLE OF INVENTION:  TARGET GENES
;    NUMBER OF SEQUENCES:  111
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/265,315
;      FILING DATE:  March 9, 1999
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/714,918
;      FILING DATE:  September 13, 1996
;      APPLICATION NUMBER:  60/009,102
;      FILING DATE:  December 22, 1995
;      APPLICATION NUMBER:  60/003,798
;      FILING DATE:  September 15, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  240/247
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  534 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-265-315-28


  Query Match             2.0%;  Score 31.2;  DB 4;  Length 534;
  Best Local Similarity   50.3%;  Pred. No. 1.5;
  Matches   72;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy   1387 atattcaatgtcattccgtattctgcacggctccacattttctgggaatgggtcgggata 1446
          |||   || ||  | || |  | |   ||||||  |  | | |       ||| | | |
Db     30 ATAACTAAGGTGGTGCCATTATGTNAGCGGCTCGTCCATGTNTATTTGGCGGTAGTGGTT 89

Qy   1447 atatccgagtattttcacagcgcgtttggaattgttttgttcgttctgctcgtcctgaag 1506
            |||  |||  | |||| |         |||      | | ||||  | | || ||||
Db     90 TAATCACTGTAGCTACACATCCAACAAATCATTCAGCATTACATTCTCGTNGCCCAGAAG 149

Qy   1507 cgcagcacatggactctaatgat 1529
          ||  |   || ||   |||||||
Db    150 CGATGGTTATTGATATTAATGAT 172


RESULT  12
US-09-265-315-28
```

```
; Sequence 28, Application US/09265315
; Patent No. 6187541
;  GENERAL INFORMATION:
;    APPLICANT:  Benton, Bret
;    APPLICANT:  Lee, Ving J.
;    APPLICANT:  Malouin, Francois
;    APPLICANT:  Martin, Patrick K.
;    APPLICANT:  Schmid, Molly B.
;    APPLICANT:  Sun, Dongxu
;    TITLE OF INVENTION:  METHODS OF SCREENING FOR COMPOUNDS
;    TITLE OF INVENTION:  ACTIVE ON STAPHYLOCOCCUS AUREUS
;    TITLE OF INVENTION:  TARGET GENES
;    NUMBER OF SEQUENCES:  111
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/265,315
;      FILING DATE:  March 9, 1999
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/714,918
;      FILING DATE:  September 13, 1996
;      APPLICATION NUMBER:  60/009,102
;      FILING DATE:  December 22, 1995
;      APPLICATION NUMBER:  60/003,798
;      FILING DATE:  September 15, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  240/247
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  534 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-265-315-28


  Query Match             2.0%;  Score 31.2;  DB 4;  Length 534;
  Best Local Similarity   50.3%;  Pred. No. 1.5;
  Matches   72;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy   1387 atattcaatgtcattccgtattctgcacggctccacattttctgggaatgggtcgggata 1446
          |||   || ||  | || |  | |   ||||||  |  | | |       ||| | | |
Db     30 ATAACTAAGGTGGTGCCATTATGTNAGCGGCTCGTCCATGTNTATTTGGCGGTAGTGGTT 89

Qy   1447 atatccgagtattttcacagcgcgtttggaattgttttgttcgttctgctcgtcctgaag 1506
            |||   |||  | |||| |         |||      | | ||||  | | || ||||
Db     90 TAATCACTGTAGCTACACATCCAACAAATCATTCAGCATTACATTCTCGTNGCCCAGAAG 149

Qy   1507 cgcagcacatggactctaatgat 1529
          ||  |   || ||   |||||||
Db    150 CGATGGTTATTGATATTAATGAT 172


RESULT  13
```

```
US-09-266-417-28
; Sequence 28, Application US/09266417
; Patent No. 6228588
;  GENERAL INFORMATION:
;    APPLICANT:  Benton, Bret
;    APPLICANT:  Lee, Ving J.
;    APPLICANT:  Malouin, Francois
;    APPLICANT:  Martin, Patrick K.
;    APPLICANT:  Schmid, Molly B.
;    APPLICANT:  Sun, Dongxu
;    TITLE OF INVENTION:  METHODS OF SCREENING FOR COMPOUNDS
;    TITLE OF INVENTION:  ACTIVE ON STAPHYLOCOCCUS AUREUS
;    TITLE OF INVENTION:  TARGET GENES
;    NUMBER OF SEQUENCES:  111
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/266,417
;      FILING DATE:  March 9, 1999
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/714,918
;      FILING DATE:  September 13, 1996
;      APPLICATION NUMBER:  60/009,102
;      FILING DATE:  December 22, 1995
;      APPLICATION NUMBER:  60/003,798
;      FILING DATE:  September 15, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  240/248
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  534 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-266-417-28


  Query Match             2.0%;  Score 31.2;  DB 4;  Length 534;
  Best Local Similarity   50.3%;  Pred. No. 1.5;
  Matches   72;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy   1387 atattcaatgtcattccgtattctgcacggctccacattttctgggaatgggtcgggata 1446
          |||   || ||  | || |  | |   ||||||  |  | | |       ||| | | |
Db     30 ATAACTAAGGTGGTGCCATTATGTNAGCGGCTCGTCCATGTNTATTTGGCGGTAGTGGTT 89

Qy   1447 atatccgagtattttcacagcgcgtttggaattgttttgttcgttctgctcgtcctgaag 1506
            |||   |||  | |||| |         |||      | | ||||  | | || ||||
Db     90 TAATCACTGTAGCTACACATCCAACAAATCATTCAGCATTACATTCTCGTNGCCCAGAAG 149

Qy   1507 cgcagcacatggactctaatgat 1529
          ||  |   || ||   |||||||
Db    150 CGATGGTTATTGATATTAATGAT 172
```

```
RESULT  14
US-08-114-072-1
; Sequence 1, Application US/08114072
; Patent No. 5624823
;  GENERAL INFORMATION:
;    APPLICANT:  Sachs, David,
;    APPLICANT:  LeGuern, Christian,
;    APPLICANT:  Megan, Sykes,
;    APPLICANT:  Blancho, Gilles
;    TITLE OF INVENTION:  INTERLEUKIN-10 AND IMMUNOSUPPRESSION
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 STATE STREET Suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/114,072
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Myers, Paul L.
;      REGISTRATION NUMBER:  35,965
;      REFERENCE/DOCKET NUMBER:  MGP-017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1365 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-114-072-1


  Query Match             2.0%;  Score 31.2;  DB 1;  Length 1365;
  Best Local Similarity   51.4%;  Pred. No. 3;
  Matches   72;  Conservative    0;  Mismatches   68;  Indels    0;  Gaps    0;

Qy    423 ctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggttag 482
          ||  |||  | || |  |    ||| |||||  |  |  ||||||  ||    |||  |
Db    745 CTGAGATATTTATTACCTCTGATACCTCAGCTCCCATTTCTATTTATTTACCGAGCCTCT 804

Qy    483 cattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaattttgggt 542
          |  | || | || |  | | |   | |  ||   |  || ||| ||  |  |||||
Db    805 CTGTGAACTATTTAGAAGAAGAACAGTATTATACTTTTTTCAGTATTTATTATTTTCACC 864

Qy    543 taccgttgacctcttcatgg 562
          |    || | || | ||| |
Db    865 TGTGTTTAAGCTTTCCATAG 884


RESULT  15
PCT-US94-09361-1
; Sequence 1, Application PC/TUS9409361
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  INTERLEUKIN-10 AND IMMUNOSUPPRESSION
;    NUMBER OF SEQUENCES:  3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII (text)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/09361
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Myers, Paul L.
;      REGISTRATION NUMBER:  35,965
;      REFERENCE/DOCKET NUMBER:  MGP-017PC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1365 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
PCT-US94-09361-1


  Query Match             2.0%;  Score 31.2;  DB 5;  Length 1365;
  Best Local Similarity   51.4%;  Pred. No. 3;
  Matches   72;  Conservative    0;  Mismatches   68;  Indels    0;  Gaps    0;

Qy    423 ctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaagaggttag 482
          ||  |||  | || |  |    ||| |||||  |  |  ||||||  ||    |||  |
Db    745 CTGAGATATTTATTACCTCTGATACCTCAGCTCCCATTTCTATTTATTTACCGAGCCTCT 804

Qy    483 cattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaattttgggt 542
          |  | || | || |  | | |   | |  ||   |  || ||| ||  |  |||||
Db    805 CTGTGAACTATTTAGAAGAAGAACAGTATTATACTTTTTTCAGTATTTATTATTTTCACC 864

Qy    543 taccgttgacctcttcatgg 562
          |    || | || | ||| |
Db    865 TGTGTTTAAGCTTTCCATAG 884


Search completed: October 30, 2001, 10:53:37
Job time: 1519 sec
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2001, 10:25:03 ; Search time 1192.62 Seconds
(without alignments)
12174.526 Million cell updates/sec

Title: US-09-619-049-854
Perfect score: 1536
Sequence: 1 atgaggcttccttgggtgat.....ggactctaatgatggattct 1536

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est75:*
137: gb_est76:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est67:*
161: gb_est68:*
162: gb_est69:*
163: gb_est70:*
164: gb_est71:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est91:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthum29:*
177: em_esthum30:*
178: em_esthum31:*
179: em_esthum32:*
180: em_esthum33:*
181: em_estom3:*
182: em_estpl11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*
190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	106.4	6.9	1027	219	CNS00KNP	AL078208 Drosophil
2	82.8	5.4	600	21	AI533834	AI533834 SD05804.5
3	82.8	5.4	686	18	AI260064	AI260064 LP03743.5
4	82.8	5.4	690	18	AI257684	AI257684 LP06122.5
5	82.8	5.4	747	4	AA264690	AA264690 LD08316.5
6	70.6	4.6	543	10	AA695295	AA695295 GM02553.5
c 7	46.8	3.0	451	122	AW944439	AW944439 SD05804.3
8	40.2	2.6	821	219	CNS00E46	AL068719 Drosophil
9	39.6	2.6	1101	219	CNS003B8	AL064086 Drosophil
10	39.4	2.6	712	239	AZ192264	AZ192264 SP_1020_B
11	38.4	2.5	1101	219	CNS00KK2	AL077673 Drosophil
c 12	37.8	2.5	445	174	BG161276	BG161276 dab68b01.
c 13	37.6	2.4	618	235	AQ878581	AQ878581 HS_3107_A
c 14	37.6	2.4	637	221	CNS03E8Y	AL240091 Tetraodon
c 15	37.2	2.4	271	187	R22732	R22732 yh19b09.r1
c 16	37.2	2.4	518	155	BG579480	BG579480 dab75e06.
c 17	37.2	2.4	621	230	AQ563399	AQ563399 HS_5335_B
18	36.8	2.4	406	150	BF542566	BF542566 UI-R-C3-s
19	36.8	2.4	448	239	AZ150028	AZ150028 SP_0005_B
c 20	36.8	2.4	458	233	AQ774098	AQ774098 HS_2269_B
21	36.8	2.4	523	118	AW639405	AW639405 b182e02.w
c 22	36.8	2.4	546	106	AU003802	AU003802 AU003802
c 23	36.8	2.4	640	106	AU005273	AU005273 AU005273
c 24	36.8	2.4	710	119	AW687502	AW687502 NF010C08R
c 25	36.8	2.4	749	239	AZ201547	AZ201547 SP_0053_A
26	36.8	2.4	752	239	AZ195546	AZ195546 SP_1030_A
27	36.6	2.4	496	22	AI634885	AI634885 tz74a06.x
28	36.6	2.4	583	32	AV719544	AV719544 AV719544
29	36.2	2.4	881	247	AZ680918	AZ680918 ENTGV56TF
c 30	36.2	2.4	976	219	CNS010DV	AL098845 Drosophil
c 31	36.2	2.4	1039	164	BE215494	BE215494 HV_CEb000
32	36	2.3	473	231	AQ620238	AQ620238 HS_5182_B
c 33	36	2.3	487	19	AI373607	AI373607 qz53b04.x
34	36	2.3	741	227	AQ365465	AQ365465 nbxb0063H
c 35	36	2.3	788	220	CNS01V5Z	AL168704 Tetraodon
c 36	36	2.3	794	235	AQ935276	AQ935276 CpG2269B
c 37	35.8	2.3	310	187	R08607	R08607 yf24c01.s1
38	35.8	2.3	499	239	AZ143021	AZ143021 SP_0041_B
39	35.6	2.3	296	25	AV157631	AV157631 AV157631
40	35.6	2.3	300	155	C07783	C07783 C07783 Yuji
41	35.6	2.3	540	233	AQ804140	AQ804140 HS_3199_A
c 42	35.6	2.3	700	106	AU002840	AU002840 AU002840
43	35.6	2.3	926	219	CNS008LM	AL052030 Drosophil
44	35.4	2.3	231	29	AV344918	AV344918 AV344918
45	35.4	2.3	298	127	BB190127	BB190127 BB190127

ALIGNMENTS

RESULT 1
CNS00KNP
LOCUS CNS00KNP 1027 bp DNA GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC: BACR18A12 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL078208
VERSION AL078208.1 GI:4957382
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1027)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES Location/Qualifiers
source 1. .1027
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR18A12"
/note="end : T7"
BASE COUNT 311 a 170 c 184 g 293 t 69 others
ORIGIN

Query Match 6.9%; Score 106.4; DB 219; Length 1027;
Best Local Similarity 53.1%; Pred. No. 1.4e-20;
Matches 191; Conservative 24; Mismatches 144; Indels 1; Gaps 1;

```
Qy       4 aggcttccttgggtgattttctgtacagttctgcttttaatttttacgaacaactcaaat 63
           ||  | | |  |   |||||   || | ||||  | ||  |  || | | | | ||  |
Db     568 AGAATGCTTCTGAATATTTTAGCTATAATTCTTGTATTCGTAATTTCAAGCCAATCGGAG 627

Qy      64 gcagatattcccgggtgcaactactacgacacggttgatatctcatacattga-aaggca 122
           |||| |||||| || ||  | |||| |||||| || ||||| ||  | ||    || ||
Db     628 GCAGTTATTCCTGGATGTGATTACTTCGACACCGTGGATATATCTCATATCCCCAAAGCT 687

Qy     123 aaacgattcgtatttatatgatgacatcgaaattcctgctagcttaactggatactacga 182
           ||| || || |||   || || ||  |    |||||:|||    | || ||    ||
Db     688 AAATGACTCCTATGCGTACGAGGAGCTAATCATTCCYGCTCATCTCACGGGTCTATATAC 747

Qy     183 gttcaggcagtttggggacggttcgattacgccgattgaaaagcatttaagggcctgtgt 242
           || ||||:  | | ||| || ||      |||  |: | |  ||  |||| ||:||| |
Db     748 ATTTAGGCMACTGGCGGATGGATCACAGGAGCCMGTKAAGAGTCACCTAAGAGCMTGTAT 807

Qy     243 ttgcagcgtgcggccctgtattcgaatctgctgcccagccaagaacttttggccaacgg 302
           :||:        || |||||:|:::||:||:|||| :|:| :|  || | || :|  |
Db     808 CYGCMAACCAAACCCTTGTATYCSRWTCYGCYGCCCCCSCMACMAAATTRTTGCGVATWG 867

Qy     303 aaaatgcgatgatggtctcaaagaggagctcgcccggttcaaaccctatatatacttcac 362
             ||    | ||| |||: |||| | :|:   :|| | || ||| :  |  |  :::|
Db     868 TYGCTGTWGGGTTGGGCTCMCAGAGAATYTBAAGBGGGTTAATCCCCMCTTTAARRWYHC 927
```

RESULT 2
AI533834
LOCUS AI533834 600 bp mRNA EST 18-MAR-1999
DEFINITION SD05804.5prime SD Drosophila melanogaster Schneider L2 cell culture pOT2 Drosophila melanogaster cDNA clone SD05804 5prime, mRNA sequence.
ACCESSION AI533834
VERSION AI533834.1 GI:4447969
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 600)

```
  AUTHORS   Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C., Brokstein
            ,P., Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Harvey, D.
            G. M. Rubin-Molecular and Cell Biology
            University of California Berkeley
            539 LSA, Berkeley, CA 94720-3200, USA
            Fax: 510 643 9947
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 58  row: A  column: 4
            High quality sequence stop: 479.
FEATURES             Location/Qualifiers
     source          1. .600
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="SD05804"
                     /clone_lib="SD Drosophila melanogaster Schneider L2 cell
                     culture pOT2"
                     /lab_host="DH5-alpha"
                     /note="Vector: pOT2; Site_1: EcoRI; Site_2: Xho1; Sized
                     fractionated cDNAs were directly ligated into pOT2.
                     Plasmid cDNA library."
BASE COUNT      167 a    134 c    140 g    159 t
ORIGIN


  Query Match             5.4%;  Score 82.8;  DB 21;  Length 600;
  Best Local Similarity   57.0%;  Pred. No. 1.4e-13;
  Matches  172;  Conservative    0;  Mismatches  127;  Indels    3;  Gaps    1;

Qy      32 ttctgcttttaatttttacgaacaactcaaatgcagatattcccgggtgcaactactacg 91
           | ||  || |  |  |||   | || ||| ||||||| |||| ||  ||| | || | ||
Db     137 TACTTGTTGTGCTGGTTATTCAAAAATCATATGCAGACATTCTCGAATGCGATTATTTCG 196

Qy      92 acacggttgatatctcatacattgaaaggcaaaacgat---tcgtatttatatgatgaca 148
           |||| || ||||| ||      | |||  | | |  ||   ||||| |||| ||| |
Db     197 ACACTGTCGATATTTCGGCAGCTCAAAAACTACAGAATGGATCGTACTTATTTGAGGGGT 256

Qy     149 tcgaaattcctgctagcttaactggatactacgagttcaggcagtttggggacggttcga 208
           |     |||| || |  || || ||| | || || || |||    |  |||||  ||||
Db     257 TGATCGTTCCGGCCATTTTGACGGGAGAATATGACTTTAGGATTCTCCCGGACGACTCGA 316

Qy     209 ttacgccgattgaaaagcatttaagggcctgtgtttgcagcgtgcggccctgtattcgaa 268
           |  | | |     ||  |||| |  ||||| ||||   ||  |||||||  |  ||
Db     317 AGCAGAAGGTGGCTCGTCACATAAGAGGATGTGTGTGCAAGCTGAAGCCCTGTGTCAGAT 376

Qy     269 tctgctgcccagccaagaacttttggccaacggaaaatgcgatgatggtctcaaagagg 328
           | ||||||||  | |  |  |  |||  || ||    ||| | ||     |    || |
Db     377 TTTGCTGCCCTCACGACCATATCATGGATAATGGTGTCTGCTACGACAACATGTCCGACG 436

Qy     329 ag 330
           ||
Db     437 AG 438


RESULT   3
AI260064
LOCUS       AI260064      686 bp    mRNA            EST       17-NOV-1998
DEFINITION  LP03743.5prime LP Drosophila melanogaster larval-early pupal pOT2
            Drosophila melanogaster cDNA clone LP03743 5prime, mRNA sequence.
ACCESSION   AI260064
VERSION     AI260064.1  GI:3867589
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 686)
  AUTHORS   Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C., Brokstein
            ,P., Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Harvey, D.
            G. M. Rubin-Molecular and Cell Biology
            University of California Berkeley
            539 LSA, Berkeley, CA 94720-3200, USA
            Fax: 510 643 9947
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 37  row: D  column: 7
            High quality sequence stop: 566.
FEATURES             Location/Qualifiers
     source          1. .686
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="LP03743"
                     /clone_lib="LP Drosophila melanogaster larval-early pupal
                     pOT2"
                     /sex="male and female"
                     /dev_stage="larvae-pupae"
                     /lab_host="DH5-alpha"
                     /note="Organ: whole body; Vector: pOT2; Site_1: EcoRI;
                     Site_2: Xho1; Sized fractionated cDNAs were directly
                     ligated into pOT2. Plasmid cDNA library.
                     "
BASE COUNT      184 a    157 c    160 g    184 t      1 others
ORIGIN


  Query Match             5.4%;  Score 82.8;  DB 18;  Length 686;
  Best Local Similarity   57.0%;  Pred. No. 1.4e-13;
  Matches  172;  Conservative    0;  Mismatches  127;  Indels    3;  Gaps    1;

Qy      32 ttctgcttttaatttttacgaacaactcaaatgcagatattcccgggtgcaactactacg 91
           | ||  || |  |  |||   | || ||| ||||||| |||| ||  ||| | || | ||
Db     120 TACTTGTTGTGCTGGTTATTCAAAAATCATATGCAGACATTCTCGAATGCGATTATTTCG 179

Qy      92 acacggttgatatctcatacattgaaaggcaaaacgat---tcgtatttatatgatgaca 148
           |||| || ||||| ||      | |||  | | |  ||   ||||| |||| ||| |
Db     180 ACACTGTCGATATTTCGGCAGCTCAAAAACTACAGAATGGATCGTACTTATTTGAGGGGT 239

Qy     149 tcgaaattcctgctagcttaactggatactacgagttcaggcagtttggggacggttcga 208
           |     |||| || |  || || ||| | || || || |||    |  |||||  ||||
Db     240 TGCTCGTTCCGGCCATTTTGACGGGAGAATATGACTTTAGGATTCTCCCGGACGACTCGA 299

Qy     209 ttacgccgattgaaaagcatttaagggcctgtgtttgcagcgtgcggccctgtattcgaa 268
           |  | | |     ||  |||| |  ||||| ||||   ||  |||||||  |  ||
Db     300 AGCAGAAGGTGGCTCGTCACATAAGAGGATGTGTGTGCAAGCTGAAGCCCTGTGTCAGAT 359

Qy     269 tctgctgcccagccaagaacttttggccaacggaaaatgcgatgatggtctcaaagagg 328
           | ||||||||  | |  |  |  |||  || ||    ||| | ||     |    || |
Db     360 TTTGCTGCCCTCACGACCATATCATGGATAATGGTGTCTGCTACGACAACATGTCCGACG 419

Qy     329 ag 330
           ||
Db     420 AG 421


RESULT   4
AI257684
LOCUS       AI257684      690 bp    mRNA            EST       17-NOV-1998
DEFINITION  LP06122.5prime LP Drosophila melanogaster larval-early pupal pOT2
            Drosophila melanogaster cDNA clone LP06122 5prime, mRNA sequence.
ACCESSION   AI257684
VERSION     AI257684.1  GI:3865209
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 690)
  AUTHORS   Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C., Brokstein
            ,P., Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
```

```
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Harvey, D.
            G. M. Rubin-Molecular and Cell Biology
            University of California Berkeley
            539 LSA, Berkeley, CA 94720-3200, USA
            Fax: 510 643 9947
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 61  row: B  column: 10
            High quality sequence stop: 592.
FEATURES             Location/Qualifiers
     source          1. .690
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="LP06122"
                     /clone_lib="LP Drosophila melanogaster larval-early pupal
                     pOT2"
                     /sex="male and female"
                     /dev_stage="larvae-pupae"
                     /lab_host="DH5-alpha"
                     /note="Organ: whole body; Vector: pOT2; Site_1: EcoRI;
                     Site_2: XhoI; Sized fractionated cDNAs were directly
                     ligated into pOT2. Plasmid cDNA library.
                     "
BASE COUNT      185 a    158 c    160 g    187 t
ORIGIN


  Query Match             5.4%;  Score 82.8;  DB 18;  Length 690;
  Best Local Similarity   57.0%;  Pred. No. 1.4e-13;
  Matches  172;  Conservative    0;  Mismatches  127;  Indels    3;  Gaps    1;

Qy      32 ttctgcttttaatttttacgaacaactcaaatgcagatattcccgggtgcaactactacg 91
           | ||  || |  |  |||   | || ||| ||||||| |||| ||  ||| | || | ||
Db     120 TACTTGTTGTGCTGGTTATTCAAAAATCATATGCAGACATTCTCGAATGCGATTATTTCG 179

Qy      92 acacggttgatatctcatacattgaaaggcaaaacgat---tcgtatttatatgatgaca 148
           |||| || ||||| ||       | |||  | | |  ||   ||||| |||| ||| |
Db     180 ACACTGTCGATATTTCGGCAGCTCAAAAACTACAGAATGGATCGTACTTATTTGAGGGGT 239

Qy     149 tcgaaattcctgctagcttaactggatactacgagttcaggcagtttggggacggttcga 208
           |    |||| || |  || || ||| | || || || |||    |  |||||  ||||
Db     240 TGCTCGTTCCGGCCATTTTGACGGGAGAATATGACTTTAGGATTCTCCCGGACGACTCGA 299

Qy     209 ttacgccgattgaaaagcatttaagggcctgtgtttgcagcgtgcggccctgtattcgaa 268
             | | | |     ||  |||| |  ||||| |||   ||  |||||||  |  ||
Db     300 AGCAGAAGGTGGCTCGTCACATAAGAGGATGTGTGTGCAAGCTGAAGCCCTGTGTCAGAT 359

Qy     269 tctgctgcccagccaagaacttttggccaacggaaaatgcgatgatggtctcaaagagg 328
           | ||||||||  | |  |  |  |||  || ||   ||| | ||     |    || |
Db     360 TTTGCTGCCCTCACGACCATATCATGGATAATGGTGTCTGCTACGACAACATGTCCGACG 419

Qy     329 ag 330
           ||
Db     420 AG 421


RESULT   5
AA264690
LOCUS       AA264690      747 bp    mRNA            EST       29-NOV-1998
DEFINITION  LD08316.5prime LD Drosophila melanogaster embryo BlueScript
            Drosophila melanogaster cDNA clone LD08316 5prime, mRNA sequence.
ACCESSION   AA264690
VERSION     AA264690.1  GI:1900714
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 747)
  AUTHORS   Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C., Brokstein
            ,P., Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Harvey, D.
            G. M. Rubin-Molecular and Cell Biology
            University of California Berkeley
            539 LSA, Berkeley, CA 94720-3200, USA
            Fax: 510 643 9947
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 83  row: B  column: 4
            High quality sequence stop: 709.
FEATURES             Location/Qualifiers
     source          1. .747
                     /organism="Drosophila melanogaster"
                     /db_xref="BDGP_EST:BDcln007659"
                     /db_xref="taxon:7227"
                     /clone="LD08316"
                     /clone_lib="LD Drosophila melanogaster embryo BlueScript"
                     /sex="male and female"
                     /dev_stage="0 to 24 hours mixed stage embryonic"
                     /lab_host="SOLR"
                     /note="Organ: embryo; Vector: BlueScript SK; Site_1: EcoRI
                     ; Site_2: XhoI; Constructed using Stratagene ZAP-cDNA
                     Synthesis kit. Oligo dT-primed and directionally cloned at
                     EcoRI and XhoI in BlueScript SK(+/-)"
BASE COUNT      202 a    169 c    171 g    205 t
ORIGIN


  Query Match             5.4%;  Score 82.8;  DB 4;  Length 747;
  Best Local Similarity   57.0%;  Pred. No. 1.5e-13;
  Matches  172;  Conservative    0;  Mismatches  127;  Indels    3;  Gaps    1;

Qy      32 ttctgcttttaatttttacgaacaactcaaatgcagatattcccgggtgcaactactacg 91
           | ||  || |  |  |||   | || ||| ||||||| |||| ||  ||| | || | ||
Db     146 TACTTGTTGTGCTGGTTATTCAAAAATCATATGCAGACATTCTCGAATGCGATTATTTCG 205

Qy      92 acacggttgatatctcatacattgaaaggcaaaacgat---tcgtatttatatgatgaca 148
           |||| || ||||| ||       | |||  | | |  ||   ||||| |||| ||| |
Db     206 ACACTGTCGATATTTCGGCAGCTCAAAAACTACAGAATGGATCGTACTTATTTGAGGGGT 265

Qy     149 tcgaaattcctgctagcttaactggatactacgagttcaggcagtttggggacggttcga 208
           |    |||| || |  || || ||| | || || || |||    |  |||||  ||||
Db     266 TGCTCGTTCCGGCCATTTTGACGGGAGAATATGACTTTAGGATTCTCCCGGACGACTCGA 325

Qy     209 ttacgccgattgaaaagcatttaagggcctgtgtttgcagcgtgcggccctgtattcgaa 268
             | | | |     ||  |||| |  ||||| |||   ||  |||||||  |  ||
Db     326 AGCAGAAGGTGGCTCGTCACATAAGAGGATGTGTGTGCAAGCTGAAGCCCTGTGTCAGAT 385

Qy     269 tctgctgcccagccaagaacttttggccaacggaaaatgcgatgatggtctcaaagagg 328
           | ||||||||  | |  |  |  |||  || ||   ||| | ||     |    || |
Db     386 TTTGCTGCCCTCACGACCATATCATGGATAATGGTGTCTGCTACGACAACATGTCCGACG 445

Qy     329 ag 330
           ||
Db     446 AG 447


RESULT   6
AA695295
LOCUS       AA695295      543 bp    mRNA            EST       28-NOV-1998
DEFINITION  GM02553.5prime GM Drosophila melanogaster ovary BlueScript
            Drosophila melanogaster cDNA clone GM02553 5prime, mRNA sequence.
ACCESSION   AA695295
VERSION     AA695295.1  GI:2697915
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 543)
  AUTHORS   Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C., Brokstein
            ,P., Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
```

```
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Harvey, D.
            G. M. Rubin-Molecular and Cell Biology
            University of California Berkeley
            539 LSA, Berkeley, CA 94720-3200, USA
            Fax: 510 643 9947
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 25  row: E  column: 5
            High quality sequence stop: 329.
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GM02553"
                     /clone_lib="GM Drosophila melanogaster ovary BlueScript"
                     /sex="female"
                     /dev_stage="newly eclosed females: germarium-stage 6"
                     /lab_host="SOLR"
                     /note="Organ: ovary; Vector: BlueScript SK; Site_1: EcoRI;
                     Site_2: XhoI; Constructed using Stratagene ZAP-cDNA
                     Synthesis kit. Oligo dT-primed and directionally cloned at
                     EcoRI and XhoI in BlueScript SK(+/-)"
BASE COUNT      135 a    139 c    144 g    124 t      1 others
ORIGIN


  Query Match             4.6%;  Score 70.6;  DB 10;  Length 543;
  Best Local Similarity   56.4%;  Pred. No. 6.1e-10;
  Matches  172;  Conservative    0;  Mismatches  129;  Indels    4;  Gaps    2;

Qy      28 acagttctgcttttaatttttacgaacaactcaaatgcagatattcccgggtgcaactac 87
           ||  |   | |||||  | ||   | | ||||||||||| || |||||||| ||| ||| |
Db      34 ACCGCATTTCTTTTGCTGTTATTGCAAAACTCAAATGCCGAAATTCCCGGTTGCGACTTC 93

Qy      88 tacgacacggttgatatctcatacattgaaaggcaa---aacgattcgtatttatatgat 144
           | |||||| || ||||| ||| |       |||       ||||  |||||  | || ||
Db      94 TTCGACACCGTAGATATTTCAAAAGCGCCAAGATTCTCGAACGGATCGTACCTCTACGAA 153

Qy     145 gacatcgaaattcctgctagcttaactggatactacgagttcaggcagtttggggacggt 204
           | | |    || || ||    || || |   | || || | || ||   | |  |||| |
Db     154 GGCTTGCTGATCCCCGCCCATTTGACAGCTGAATATGACTACAAGCTCCTGGCCGACGAT 213

Qy     205 tcgattacgccgattgaaaagcatttaagggcctgtgtttgcagcgtgcggccctgtatt 264
           ||||    |  |  ||   |||   || | |  |||| |||  | |  |||| || |||
Db     214 TCGAAGGAGAAG-GTGGCGAGCCAGTACGAGGATGTGCCTGCCACCTCAGGCCATGCATT 272

Qy     265 cgaatctgctgcccagccaagaacttttggccaacggaaaatgcgatgatggtctcaaa 324
           ||  | || |||||      | |      ||   ||  | || ||| | |  |   |
Db     273 CGGTTTTGTTGCCCCCAGTACCAAAAGATGCAAAAGAGCAAGTGCTACGGCGACATGTCG 332

Qy     325 gagga 329
           |||||
Db     333 GAGGA 337


RESULT   7
AW944439/c
LOCUS       AW944439      451 bp    mRNA            EST       31-MAY-2000
DEFINITION  SD05804.3prime SD Drosophila melanogaster Schneider L2 cell culture
            pOT2 Drosophila melanogaster cDNA clone SD05804 3 similar to
            AF109308: Drosophila melanogaster putative G-protein-coupled
            receptor (mth) mRNA, complete cds, mRNA sequence.
ACCESSION   AW944439
VERSION     AW944439.1  GI:8122170
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 451)
  AUTHORS   Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C., Brokstein
            ,P., Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: SD05804.5prime
            Contact: Harvey, D.
            G. M. Rubin-Molecular and Cell Biology
            University of California Berkeley
            539 LSA, Berkeley, CA 94720-3200, USA
            Fax: 510 643 9947
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Based upon the presence of a XhoI site followed by a run of 14 or
            more T residues at the beginning of the sequence, this clone was
            polyadenylated. The resulting Poly-T sequence has been removed.
            Plate: SD.58  row: A  column: 4
            High quality sequence stop: 450.
FEATURES             Location/Qualifiers
     source          1. .451
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="SD05804"
                     /clone_lib="SD Drosophila melanogaster Schneider L2 cell
                     culture pOT2"
                     /lab_host="DH5-alpha"
                     /note="Vector: pOT2; Site_1: EcoRI; Site_2: Xho1; Sized
                     fractionated cDNAs were directly ligated into pOT2.
                     Plasmid cDNA library."
BASE COUNT      142 a     79 c     87 g    143 t
ORIGIN


  Query Match             3.0%;  Score 46.8;  DB 122;  Length 451;
  Best Local Similarity   54.7%;  Pred. No. 0.0079;
  Matches   93;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy    1363 tccatcgtgatgggccttacttggatattcaatgtcattccgtattctgcacggctccac 1422
           | |||| | |||||  |  | ||||  ||  |  |      | || | | | |    | ||
Db     448 TTCATCATTATGGGATTGTCCTGGAGCTTGGAGATAGGCTCCTACTTTTCGCAATCCAAC 389

Qy    1423 attttctgggaatgggtcgggataatatccgagtattttcacagcgcgtttggaattgtt 1482
             ||||     |||    |  |  | || |||||  |  |  |    |||||  |
Db     388 CAAACTTGGGCCAACGTCTTTTTGGTGGCTGACTATTTAAATTGGTCTCAAGGAATCATC 329

Qy    1483 ttgttcgttctgctcgtcctgaagcgcagcacatggactctaatgatgga 1532
           | ||  | ||| |||| |||||||||||||| ||||  ||  ||  |||
Db     328 ATATTTATACTGTTCGTTCTGAAGCGCAGCACGTGGAGACTCTTGCAGGA 279


RESULT   8
CNS00E46
LOCUS       CNS00E46      821 bp    DNA             GSS       04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR28O12 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL068719
VERSION     AL068719.1  GI:4948862
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 821)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
```

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES             Location/Qualifiers
     source          1. .821
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR28O12"
                     /note="end : T7"
BASE COUNT      214 a    183 c    129 g    206 t     89 others
ORIGIN
```

```
  Query Match             2.6%;  Score 40.2;  DB 219;  Length 821;
  Best Local Similarity   44.6%;  Pred. No. 0.91;
  Matches   78;  Conservative   17;  Mismatches   80;  Indels    0;  Gaps    0;

Qy     677 taattataacaatgatatgctttgtcctaacaatcgcagtatatctatacattaagaagc 736
           || | |||:|  ||| | ||    ::|:||: | :   || || || ||  |  ||||:|
Db     229 TAGTGATAWCCTTGAAAAGCAAAAWMCWAAMGAAMARCGTGTACCTCTATGTCGAGAARC 288

Qy     737 tgcgaaatgttactggcaagtgcattgtatgctgtatagtttcgaggtttatccagtgct 796
           | || ||  :    || ||:||  |  : :||:   | |  ||   |||  |    |  |
Db     289 TTCGCAACCYGCACGGAAARTGTTTCAWCYGCYACTTGGCCTCTTTGTTCTTGGGCTATT 348

Qy     797 tgatcatgatactagatcatttaaatctattgaatggcatttgctctccagctgg 851
           |  || :| | ||  |    |  ||   :|     ||| | |||  | :|:| :|
Db     349 TCTTCCYGGTCCTCAACGTGTGGAAAWAMTCCCCCGGCTTCTGCGTTAMARCARG 403
```

```
RESULT   9
CNS003B8
LOCUS        CNS003B8      1101 bp     DNA              GSS        03-JUN-1999
DEFINITION   Drosophila melanogaster genome survey sequence TET3 end of BAC #
             BACR08I08 of RPCI-98 library from Drosophila melanogaster (fruit
             fly), genomic survey sequence.
ACCESSION    AL064086
VERSION      AL064086.1  GI:4941842
KEYWORDS     GSS.
SOURCE       fruit fly.
  ORGANISM   Drosophila melanogaster
             Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
             Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
             Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE    1  (bases 1 to 1101)
  AUTHORS    Genoscope.
  TITLE      Direct Submission
  JOURNAL    Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
             BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
             - Web : www.genoscope.cns.fr)
COMMENT      Determination of this BAC-end sequence was carried out as part of a
             collaboration with the Berkeley Drosophila Genome Project (BDGP).
             The BDGP is constructing a physical map of the Drosophila
             melanogaster genome using these BACs. For further information
             please see  http://www.fruitfly.org The BDGP Drosophila
             melanogaster BAC library was prepared by Kazutoyo Osoegawa and
             Aaron Mammoser in Pieter de Jong's laboratory in the Department of
             Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
             NY. The library is named RPCI-98 and was constructed by partial
             EcoRI digestion of Drosophila DNA provided by the BDGP from the
             isogenic strain y2; cn bw sp, the same strain used for the BDGP's
             P1 and EST libraries. A more detailed description of the library
             and how to order individual BAC clones, the entire library, or
             filters for hybridization from the BACPAC Resource Center can be
             found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR08I08"
                     /note="end : T3"
BASE COUNT      271 a    101 c    296 g    214 t     219 others
ORIGIN
```

```
  Query Match             2.6%;  Score 39.6;  DB 219;  Length 1101;
  Best Local Similarity   56.2%;  Pred. No. 1.5;
  Matches  113;  Conservative    1;  Mismatches   81;  Indels    6;  Gaps    2;

Qy    1338 tagttacctacagttcctgcggctctccatcgtgatgggccttacttggatattcaatgt 1397
           ||||||| ||:| ||| |||| || |   |  | ||||||   | |||| || |  |
Db      28 TAGTTACATAMAATTCGTGCGACTTTWTCTAATTATGGGCGCAAGTTGGCTACTTGACCA 87

Qy    1398 cattccgtattctgcacgg----ctccacattttctgggaa--tgggtcgggataatatc 1451
             |  ||   |  ||   |    ||| | ||| ||||  |     |    ||   || |
Db      88 ATTAACGCGATTGGCTGAGGATTCTCACCTTTTGCTGGACACTATAGNTCTGAACTTAAC 147

Qy    1452 cgagtattttcacagcgcgtttggaattgttttgttcgttctgctcgtcctgaagcgcag 1511
             |  || ||  |  ||| ||||||||  |  | || ||  ||||  | |||||| ||||
Db     148 AGTCTACTTAAATGCCGCCTTTGGAATACTCATATTTGTSTTGCTTATACTGAAGGGCAG 207

Qy    1512 cacatggactctaatgatgga 1532
           |||||      | || |||||
Db     208 CACATTTCASATGATAATGGA 228
```

```
RESULT  10
AZ192264
LOCUS        AZ192264      712 bp    DNA              GSS        30-AUG-2000
DEFINITION   SP_1020_B2_H06_T7A Strongylocentrotus purpuratus, purple sea urchin
             , sperm genomic BAC library Strongylocentrotus purpuratus genomic
             clone Plate=1020 Col=12 Row=P, DNA sequence.
ACCESSION    AZ192264
VERSION      AZ192264.1  GI:8375443
KEYWORDS     GSS.
SOURCE       Strongylocentrotus purpuratus.
  ORGANISM   Strongylocentrotus purpuratus
             Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
             Echinoidea; Euechinoidea; Echinacea; Echinoida;
             Strongylocentrotidae; Strongylocentrotus.
REFERENCE    1  (bases 1 to 712)
  AUTHORS    Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,
             Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray
             ,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J, Davidson,E.H. and
             Hood,L.
  TITLE      A sea urchin genome project: Sequence scan, virtual map, and
             additional resources
  JOURNAL    Proc. Natl. Acad. Sci. U. S. A. 97 (17), 9514-9518 (2000)
  MEDLINE    20402566
COMMENT      Contact: Cameron, RA, Davidson, EH, Hood, L
             Division of Biology 156-29
             California Institute of Technology
             Pasadena California 91125, USA
             Tel: (626) 395-8421
             Fax: (626) 793-3047
             Email: acameron@caltech.edu
             Plate: 1020  row: P  column: 12
             Seq primer: T7
             Class: BAC ends
             High quality sequence stop: 712.
FEATURES             Location/Qualifiers
     source          1. .712
                     /organism="Strongylocentrotus purpuratus"
                     /db_xref="taxon:7668"
                     /clone="Plate=1020 Col=12 Row=P"
                     /clone_lib="Strongylocentrotus purpuratus, purple sea
                     urchin, sperm genomic BAC library"
                     /note="Organ: sperm; Vector: BACe3.6; BAC Clones in E-Coli
                     DH10B"
```

```
BASE COUNT      196 a    129 c    138 g    249 t
ORIGIN

  Query Match             2.6%;  Score 39.4;  DB 239;  Length 712;
  Best Local Similarity   50.3%;  Pred. No. 1.5;
  Matches   97;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;

Qy     534 attttgggttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaa 593
           | ||||||| ||  |||||| || |||| |||    |  || ||  |  | | | || ||
Db      70 ACTTTGGGTAACGCTTGACCCCTACATGAAAACGTGGACCTGCTTTTCTTCCTGTCAAAA 129

Qy     594 ctttgattcggattttccaaagtccatgtggataatacgacaccgctgcacaagccacat 653
           | | || |    | || ||  ||  || | |||||        |   ||| | | |   |
Db     130 CGTGGACTTCTTTATTTCAGGGTGTATTTTGATAACTGTTGTACTTGGCAAATGACTAGT 189

Qy     654 atctcctggatccttagagattctaattataacaatgatatgctttgtcctaacaatcgc 713
           |     |     ||| | | |  | ||  || ||  |     ||  |  ||  ||
Db     190 ACAAATTTCTCGTTTAAAAACTTCATTTTGAAAAAAAAAGGAATTCCTTTTAGTAACATG 249

Qy     714 agtatatctatac 726
           || |||| | |
Db     250 TTTAAATCTCTTC 262


RESULT  11
CNS00KK2
LOCUS       CNS00KK2     1101 bp    DNA             GSS       03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC:
            BACR17F22 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL077673
VERSION     AL077673.1  GI:4957249
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR17F22"
                     /note="end : T7"
BASE COUNT      191 a     94 c    163 g    151 t    502 others
ORIGIN

  Query Match             2.5%;  Score 38.4;  DB 219;  Length 1101;
  Best Local Similarity   13.7%;  Pred. No. 3.4;
  Matches   59;  Conservative  126;  Mismatches  245;  Indels    0;  Gaps    0;

Qy    1012 agtatttatatagttaatcagatttgggaaaacgatcccagtaaatggaactggttgcct 1071
           : |:  |          :|   :   |        |       |     ||    |   |
Db     582 MNTMNNTGNTGNNGNNMTNNMMMMMGNNNNNNNGTNNNNTGNNGNNGGGNNNGTNGGNTG 641

Qy    1072 ctggtcggttttattcggtgctcggtcaaagattggcacccatccgtctggatctatata 1131
           | :|  |      || |||::       | |  :|      || |     : :
Db     642 TTTVTNMGGNGNNTTKGGTVMMGMNNNNNTGKTCNNMANNNNGGGNTCNGNNMNNNKNKN 701

Qy    1132 agtggaccgtcgctggccctgagcactttcaatgtcgccatgttcgccctgacagccatt 1191
              :    :   :   :      :  |  |       :     :::     : :|
Db     702 NNGNKNNNNNMMNNNKNNGKNNTNNTKNNCNNTNNNMNKNKMNNNNMMMNMNKNKKMAGN 761

Qy    1192 tacattaggaaagtgaaggggggtataaataagttcacaaatgaggaggagggaaggatt 1251
           |:|:   |::|   |:     ::: | | ::    ::::: ::| :||:|:|:: |::
Db     762 MAMAKAKAGMMAMMAAMAMAMAKBDMAKAMMMMMMGMMMMMVKMGMMGGMGKGHMMGRKM 821

Qy    1252 aactgcataaactttgacagccagacgtaagatatgataaaaccttggaccttggtaaac 1311
              ::: :: :   |  ::: |:  :: :::   : : : |:|: : :: :  ::::
Db     822 GKTMMMRKNMVNKMGMAKNKMMBMAMMMMMMMKMMNKDKMNMWCKTKCMNMKMKYMMMMM 881

Qy    1312 tagttatccttgccgtattatccacatagttacctacagttcctgcggctctccatcgtg 1371
           : : : :::  :  ::: ::: :: ::|| : |:  :  : | ::   ::  :  :  :
Db     882 MMAKMMMHMKMMMKMKHKMVHMKMMCHMGTCMGCKKKMMAKTCMBMMMAKMAKVGASMAK 941

Qy    1372 atgggccttacttggatattcaatgtcattccgtattctgcacggctccacattttctgg 1431
           :: :|:   |:    :    ::| |::   ::|    : | | : :::  :|   :: |:
Db     942 DKMDGMAGMAMMGNCMGCMGMMAGGKMKMMMMGAGVAMGGTAGKMMKMGTMAMMAKMAGK 1001

Qy    1432 gaatgggtcg 1441
           .  :|:  : :|
Db    1002 MMAKHAVNMG 1011


RESULT  12
BG161276/c
LOCUS       BG161276      445 bp    mRNA            EST       06-FEB-2001
DEFINITION  dab68b01.x1 NICHD XGC Emb4 Xenopus laevis cDNA clone IMAGE:4202064
            3' similar to SW:PRS6_HUMAN P43686 26S PROTEASE REGULATORY SUBUNIT
            6B ;, mRNA sequence.
ACCESSION   BG161276
VERSION     BG161276.1  GI:12695195
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 445)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: Xenopus clones from this library are available
            through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 213.
FEATURES             Location/Qualifiers
     source          1. .445
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="IMAGE:4202064"
                     /clone_lib="NICHD XGC Emb4"
                     /dev_stage="embryo, stage 31-32"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: whole embryo; Vector: pCMV-SPORT6; Site_1:
```

NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."
BASE COUNT 112 a 98 c 114 g 121 t
ORIGIN

Query Match 2.5%; Score 37.8; DB 174; Length 445;
Best Local Similarity 49.3%; Pred. No. 4;
Matches 99; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

Qy 416 atgagttctttgattgtgatgaaatgatttacatcagcgactttaactatttcttagaag 475
Db 282 AGGACTATGTTGCTCGTCCTGATAAGATTTCCGGGGCCGACATTAACTCCATCTGCCAAG 223

Qy 476 aggttagcattcaaatcttcaataaatgtggacttatagtttggttccaggatggaaaat 535
Db 222 AGGCTGGAATGCTTGCTGTCCGGGAGAACCGATACATAGTGTTGGCCAAGGACTTTGAGA 163

Qy 536 tttgggttaccgttgacctcttcatggaaaaacaggactactgcttgtaccggcacaact 595
Db 162 AAGCCTATAAGACGGTCATCAAGAAGGATGAACAGGAGCACGAGTTCTACAAATAAACAT 103

Qy 596 ttgattcggatttccaaagt 616
Db 102 TAACTCCCTGTCCTCCCAACT 82

RESULT 13
AQ878581/c
LOCUS AQ878581 618 bp DNA GSS 09-NOV-1999
DEFINITION HS_3107_A1_H09_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3107 Col=17 Row=O, DNA sequence.
ACCESSION AQ878581
VERSION AQ878581.1 GI:6310144
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 618)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 3107 row: O column: 17
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 618.
FEATURES Location/Qualifiers
source 1. .618
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=3107 Col=17 Row=O"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"
BASE COUNT 218 a 97 c 64 g 239 t
ORIGIN

Query Match 2.4%; Score 37.6; DB 235; Length 618;
Best Local Similarity 59.3%; Pred. No. 5;
Matches 64; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 674 ttctaattataacaatgatatgctttgtcctaacaatcgcagtatatctatacattaaga 733
Db 566 TTGTAATAATATCAATGTTATTTTAGTGTCTAATAATCCCATTATAGATGACATATACAA 507

Qy 734 agctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcga 781
Db 506 ATAGGAGAAAGATCATCGTGAAGAGAAATGTATTATTAATAGTTCTGA 459

RESULT 14
CNS03E8Y/c
LOCUS CNS03E8Y 637 bp DNA GSS 17-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone 019F16 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION AL240091
VERSION AL240091.1 GI:7960860
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 637)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Charaterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 637)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 637)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.
FEATURES Location/Qualifiers
source 1. .637
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="019F16"
/clone_lib="G"
/note="Genoscope sequence ID : C0BG019DC08LP1-end : T7"
BASE COUNT 146 a 131 c 126 g 233 t 1 others
ORIGIN

Query Match 2.4%; Score 37.6; DB 221; Length 637;
Best Local Similarity 56.5%; Pred. No. 5.1;
Matches 70; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 668 tagagattctaattataacaatgatatgctttgtcctaacaatcgcagtatatctataca 727
Db 631 TAATAATAATAATAATAATAATAATAGGAGGGGTACTAACCCTCCCATTATAAATGTTTA 572

Qy 728 ttaagaagctgcgaaatgttactggcaagtgcattgtatgctgtatagtttcgaggttta 787

```
Db    571 TTCTCATTCTGCTGATTGCCACTAGAGGATGAACAAAATGCTCTGAATTCAGAAGGTGCA 512

Qy    788 tcca 791
          || |
Db    511 TCTA 508


RESULT  15
R22732/c
LOCUS       R22732         271 bp    mRNA            EST       19-APR-1995
DEFINITION  yh19b09.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:130169 5', mRNA sequence.
ACCESSION   R22732
VERSION     R22732.1  GI:777489
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 271)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1280
            High quality sequence stops: 209
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1280   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 209.
FEATURES             Location/Qualifiers
     source          1. .271
                     /organism="Homo sapiens"
                     /db_xref="GDB:536660"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:130169"
                     /clone_lib="Soares placenta Nb2HP"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.  "
BASE COUNT       82 a     36 c     59 g     92 t      2 others
ORIGIN


  Query Match             2.4%;  Score 37.2;  DB 187;  Length 271;
  Best Local Similarity   53.1%;  Pred. No. 5.2;
  Matches   78;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy    635 accgctgcacaagccacatatctcctggatccttagagattctaattataacaatgatat 694
          | | |||  ||| ||| |||| || |   ||  |||| | ||   | | |  ||   ||
Db    261 AACACTGAGCAATCCAGATATTCCCAGNCCCCGGAGAGTTACTGCCTTTCATGATCTGAT 202

Qy    695 gctttgtcctaacaatcgcagtatatctatacattaagaagctgcgaaatgttactggca 754
          |||||  | | | |   |||  | |  || |  || || | |  ||||   |

Db    201 CTTTTGTAATGAAACTACAAGTGCAACATTAAAAAAAAAAACAGATAAATAGAAACCCTG 142

Qy    755 agtgcattgtatgctgtatagtttcga 781
            |  ||| |  ||| || | ||| ||
Db    141 GTTTTATTTTCAGCTTTACATTTTTGA 115


Search completed: October 30, 2001, 10:52:27
Job time: 1644 sec
```

This Page Blank (uspto)